# A new criterion for the logarithmic Sobolev inequality and two applications

**PREPRINT of article published in**
**Journal of Functional Analysis 243 (2007) 121-157.**

Felix Otto and Maria G. Reznikoff [a,b]

[a] *Inst. für Angewandte Mathematik, Universität Bonn, Germany,*
*otto@iam.uni-bonn.de*

[b] *Mathematics Dept., Princeton University, Princeton, USA,*
*mrezniko@math.princeton.edu*

---

## Abstract

We present a criterion for the logarithmic Sobolev inequality (LSI) on the product space $X_1 \times \ldots \times X_N$. We have in mind an $N$–site lattice, unbounded continuous spin variables, and Glauber dynamics. The interactions are described by the Hamiltonian $H$ of the Gibbs measure. The criterion for LSI is formulated in terms of the LSI constants of the single–site conditional measures and the size of the off–diagonal entries of the Hessian of $H$. It is optimal for Gaussians with positive covariance matrix. To illustrate, we give two applications: one with weak interactions and one with strong interactions and a decay of correlations condition.

*Key words:* logarithmic Sobolev inequality, decay of correlations, Glauber dynamics
60K35 (primary), 60J25, 82B21

---

## 1 Introduction

The logarithmic Sobolev inequality (LSI) was introduced by Gross [G]. It is attached to a Markov semi–group $P_t$ with reversible invariant measure $\mu$. We refer to the recent survey paper [GZ, Chapters 1 - 4] for a general introduction within the framework of $\Gamma_1$–calculus. Like the spectral gap inequality (SGI), which is analytically speaking a Poincaré inequality for the measure $\mu$, LSI yields exponential convergence of the Markov semi–group to equilibrium with a rate given by the constant in the inequality. Like the classical Sobolev–Poincaré inequalities, the LSI also expresses an improved integrability. Gross made two important observations: On the one hand, this improved integrability is strong enough to yield hypercontractivity for the Markov semi–group $P_t$; see for instance [GZ, Theorem 4.1]. Hypercontractivity is a sharpened statement of the trend to equilibrium; see [GZ, Section 4.1]. On the other hand, the improved integrability is weak enough (the gain is just a logarithm) to be stable under Cartesian products (of the Markov semi–groups and their reversible invariant measures); see [GZ, Theorem 4.4] and Remark 1. These are the features that make the LSI suitable for spin systems.

There are only a few sufficient criteria for LSI. The first important criterion, due to Holley & Stroock [HS], is perturbative in nature and not ideally suited for spin systems. The second important criterion, due to Bakry & Emery [BE], is non–perturbative, but structurally quite restrictive, cf. Remark 2. The criterion of Bakry & Emery is based on the $\Gamma_2$–calculus, which essentially requires a Riemannian spin space. This is also the framework we adopt. We shall frequently refer to [L] for a nice review.

Our main result (Theorem 1) is a clean sufficient criterion for LSI. We consider a Gibbs measure $\mu$ on a product space $X_1 \times \ldots \times X_N$. We formulate the condition in terms of the Hamiltonian and the LSI constants of the single–site conditional measures. The result can be viewed as an adaptation of the above–mentioned product argument to allow for coupling, cf. Remark 1. It is indeed important to start from the product argument, since a naive application of the Holley–Stroock principle (see for instance [L, Lemma 1.2]) would yield an LSI–constant that increases exponentially with the number $N$ of sites, no matter how weak the interaction. We require weaker hypotheses than the Bakry-Emery principle (cf. Remark 2): We do not require strict convexity of the Hamiltonian. Moreover, for $X_i = \mathbb{R}$ and attractive interactions, the bound of Theorem 1 is sharp for Gaussians (cf. Remark 4). For the SGI of a Gibbs measure, a result similar to Theorem 1 and somewhat stronger is proved by Ledoux [L], cf. Remark 3.

Earlier work of Royer [R, Théorème 5.2.1] based on Zegarlinski's iterative method produces a similar, but weaker, bound for the LSI constant (cf. Remark 5). Introduced by Zegarlinski in 1990 [Z1], the iterative technique was applied and developed by Zegarlinski [Z2,Z3] and Stroock & Zegarlinski [SZ].

A second approach that has been widely used in the analysis of spin systems is the Lu–Yau martingale method (introduced in [LY] and also reviewed in [L, Section 5]). The martingale method relies on the LSI for *marginals* (i.e. averaged out versions) of the conditional measures; see for instance [LY, (6.3)], [Yo1, Lemma 3.2], [L, Proposition 4.1]. In the case of unbounded spin space, these LSI's for marginals rely in turn on global spectral gap estimates, cf. [Yo1, Theorem 2.2] and [L, Proposition 3.1]. Recent progress by Blower and Bolley

[BB, Theorem 1.3] also transforms information about LSI for conditional and marginal measures into a global LSI.

Our functional analytic approach combines the advantages of both approaches described above: It avoids the fixed point iteration and it requires the LSI for conditional measures, but not marginals. It grows out of work presented in [O,OV,GORV]. See Section 3 for additional comments regarding connections among the different methods.

We begin in Subsection 1.1 by presenting the main result. In Subsection 1.2 we state a two–scale criterion for LSI (introduced in [GORV]) which we will use in the applications. Section 2 contains the examples. In Section 3 we present some auxiliary results, and finally in Section 4 we give the proofs of these lemmas and Theorems 1 and 2.

### *1.1 Main result and discussion*

We deal only with Euclidean spaces $X$, although our arguments would also go through for general Riemannian manifolds. Norms $|\cdot|$ and the notion of gradient $\nabla$ are derived from the Euclidean structure. We will use the notation $\mathcal{P}(X)$ to denote the space of probability measures on $X$.

The logarithmic Sobolev inequality can be defined in the following way:

**Definition 1 (LSI)** *Let* $\Phi(x) := x \log x$. *The probability measure* $\mu(dx) \in \mathcal{P}(X)$ *satisfies the logarithmic Sobolev inequality LSI(*$\rho$*) with constant* $\rho$ *if*

$$\forall\, f(x) \geq 0 \quad \int \Phi(f)\, d\mu - \Phi\left(\int f\, d\mu\right) \;\leq\; \frac{1}{\rho} \int \frac{1}{2f} |\nabla f|^2\, d\mu. \tag{1}$$

We recall the disintegration of a probability measure into a conditional probability measure and the corresponding marginal:

**Definition 2 (Conditional and marginal measures)** *To any probability measure* $\mu(dx_1dx_2) \in \mathcal{P}(X_1 \times X_2)$ *we associate the marginal* $\bar{\mu}(dx_1) \in \mathcal{P}(X_1)$ *and the family of conditional measures* $\mu(dx_2|x_1) \in \mathcal{P}(X_2)$ *via*

$$\forall\, \zeta(x_1,x_2) \qquad \int \zeta(x_1,x_2)\,\mu(dx_1dx_2) \;=\; \int\int \zeta(x_1,x_2)\,\mu(dx_2|x_1)\,\bar{\mu}(dx_1).$$

Our main result is:

**Theorem 1** *Let* $X_1, \cdots, X_N$ *be Euclidean spaces and* $\mu(dx_1 \cdots dx_N)$ *a probability measure on the product space* $X_1 \times \cdots \times X_N$ *with a smooth positive Lebesgue density* $\frac{d\mu}{d\mathcal{L}}$.

*We assume that for all* $i < j \in \{1, \cdots, N\}$ *there exists* $\kappa_{ij} < \infty$ *such that the Hamiltonian* $H(x_1, \cdots, x_N) = -\log \frac{d\mu}{d\mathcal{L}}$ *satisfies*

$$\forall\, (x_1, \cdots, x_N) \quad |\nabla_i \nabla_j H(x_1, \cdots, x_N)| \;\le\; \kappa_{ij}. \tag{2}$$

*Here and in what follows,* $|\nabla_i \nabla_j H(x_1, \cdots, x_N)|$ *denotes the operator norm of the bilinear form* $\nabla_i \nabla_j H(x_1, \cdots, x_N)$ *on* $X_i \times X_j$.

*We assume that for all* $i \in \{1, \cdots, N\}$ *there exists* $\rho_i > 0$ *such that*

$$\begin{aligned} &\forall\, (x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_N) \\ &\mu(dx_i|x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_N) \quad \text{satisfies } LSI(\rho_i). \end{aligned} \tag{3}$$

*Consider the symmetric $N \times N$–matrix $A$ defined by*

$$\begin{aligned} A_{ij} &= -\kappa_{ij} \ \textit{for}\ i < j \in \{1, \cdots, N\}, \\ A_{ii} &= \ \ \rho_i \quad \textit{for} \quad i \in \{1, \cdots, N\}. \end{aligned} \tag{4}$$

*We assume that there exists $\rho > 0$ such that we have in the sense of quadratic forms*

$$A \ \geq \ \rho \,\mathrm{id}. \tag{5}$$

*Then*

$$\mu(dx_1 \cdots dx_N) \quad \textit{satisfies LSI}(\rho). \tag{6}$$

**Remark 1 (Product measures)** *Assume that we are in the setting of Theorem 1 and that in addition $\mu$ is a product measure, i.e.*

$$\mu(dx_1 \cdots dx_N) \ = \ \mu_1(dx_1) \ldots \mu_N(dx_N). \tag{7}$$

*It has been known since its origins [G, Remark 3.3] (see [L, Lemma 1.1] for a modern presentation) that the LSI is compatible with taking products. More precisely, if for each $i \in \{1, \cdots, N\}$, there exists $\rho_i > 0$ such that*

$$\mu_i(dx_i) \quad \textit{satisfies LSI}(\rho_i),$$

*then*

$$\mu(dx_1 \ldots dx_N) \quad \textit{satisfies LSI}(\rho)$$

*with*

$$\rho \ := \ \min\{\rho_1, \cdots, \rho_N\}. \tag{8}$$

*Theorem 1 matches this bound for product measures: In case of (7), we have $\nabla_i \nabla_j H \equiv 0$ for all $i < j \in \{1, \cdots, N\}$. Therefore, we may choose $\kappa_{ij} = 0$, to* *the effect that the optimal $\rho$ in (5) is precisely (8). Theorem 1 can be interpreted as a perturbation of the above product property.*

**Remark 2 (The criterion of Bakry–Emery)** *In this remark, we relate Theorem 1 to the Bakry–Emery criterion [BE]; see for instance [L, Corollary 1.6] for an efficient proof. In the notation of Theorem 1, the Bakry–Emery principle reads as follows: Consider the symmetric $N \times N$–matrix $A(x_1, \cdots, x_N)$ defined by*

$$A_{ij}(x_1, \cdots, x_N) \;=\; \nabla_i \nabla_j H(x_1, \cdots, x_N) \quad \textit{for } \; i, j \in \{1, \cdots, N\}.$$

*If $\rho > 0$ is such that*

$$\forall \, (x_1, \cdots, x_N) \quad A(x_1, \cdots, x_N) \;\geq\; \rho \, \mathrm{id} \tag{9}$$

*in the sense of quadratic forms, then*

$$\mu \quad \textit{satisfies LSI}(\rho).$$

*On the one hand, Theorem 1 is stronger than the Bakry–Emery principle, since it assumes less about the single site conditional measures,*

$$\mu(dx_i | x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_N).$$

*On the other hand, Theorem 1 is somewhat weaker, since for $i < j \in \{1, \cdots, N\}$, the criterion (5) appeals to $A_{ij} = -\sup_{(x_1, \cdots, x_N)} |\nabla_i \nabla_j H(x_1, \cdots, x_N)|$, whereas the criterion (9) involves just $\nabla_i \nabla_j H(x_1, \cdots, x_N)$.*

**Remark 3 (Spectral gap inequality)** *In this remark, we relate Theorem 1 to what is known about the spectral gap. Recall that a probability measure $\mu(dx)$ is said to satisfy the spectral gap estimate SGI($\rho$) with constant $\rho > 0$*

*provided*

$$\forall\, h(x) \quad \int h^2\,d\mu - \Big(\int h\,d\mu\Big)^2 \;\le\; \frac{1}{\rho}\int |\nabla h|^2\,d\mu. \tag{10}$$

*It is well–known that SGI($\rho$) is a consequence of LSI($\rho$), as can be seen by using $f = 1 + \epsilon h$ in (1) and expanding to second order in $\epsilon$.*

*In a situation analogous to Theorem 1, a somewhat stronger result is known on the level of spectral gap, cf. [L, Proposition 3.1]. We now state this result in the notation of Theorem 1: One assumes that for $i \in \{1, \cdots, N\}$, there exists $\rho_i > 0$ such that*

$$\begin{array}{l} \forall\,(x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_N) \\ \mu(dx_i|x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_N) \quad \textit{satisfies SGI}(\rho_i). \end{array}$$

*One considers the symmetric $N\times N$–matrix $A(x_1,\cdots,x_N)$ defined through*

$$\begin{array}{lcll} A_{ij}(x_1,\cdots,x_N) &=& \nabla_i\nabla_j H(x_1,\cdots,x_N) & \textit{for } i<j\in\{1,\cdots,N\}, \\ A_{ii}(x_1,\cdots,x_N) &=& \rho_i & \textit{for} \quad i\in\{1,\cdots,N\}, \end{array}$$

*and assumes that there exists $\rho > 0$ such that in the sense of quadratic forms*

$$\forall\,(x_1,\cdots,x_N) \quad A(x_1,\cdots,x_N) \;\ge\; \rho\,\mathrm{id}.$$

*Then one has*

$$\mu(dx_1\cdots dx_N) \quad \textit{satisfies SGI}(\rho).$$

**Remark 4 (Gaussians)** *We assume $X_i = \mathbb{R}$ for $i \in \{1,\cdots,N\}$ and let $A$ be a symmetric positive definite $N\times N$–matrix. In this remark, we argue that Theorem 1 is optimal for Gaussians, i.e. for Hamiltonians of the form:*

$$H \;=\; -\log\frac{d\mu}{d\mathcal{L}} \;=\; \frac{1}{2}\sum_{i,j\in\{1,\cdots,N\}} x_i\,A_{ij}\,x_j + \sum_{i\in\{1,\cdots,N\}} b_i\,x_i, \tag{11}$$

*provided the coupling is attractive:*

$$A_{ij} \ \leq \ 0 \quad \textit{for}\, i < j \in \{1, \cdots, N\}. \tag{12}$$

*Recall that the covariance matrix is given by the inverse $A^{-1}$ of $A$:*

$$(A^{-1})_{ij} \ = \ \int x_i\, x_j\, d\mu - \int x_i\, d\mu \int x_j\, d\mu. \tag{13}$$

*Incidentally, according to Lemma 9 below, the attractive coupling (12) implies non–negative covariances*

$$\int x_i\, x_j\, d\mu - \int x_i\, d\mu \int x_j\, d\mu \ \geq \ 0 \quad \textit{for}\, i, j \in \{1, \cdots, N\}.$$

*We also recall that for $\mu(dx)$ of the form (11) and any number $\rho > 0$, we have the equivalences*

$$\begin{aligned} & A \ \geq \ \rho\, \mathrm{id} \quad \textit{as quadratic forms} & (14)\\ & \Longleftrightarrow \mu \quad \textit{satisfies LSI}(\rho) & (15)\\ & \Longleftrightarrow \mu \quad \textit{satisfies SGI}(\rho). & (16) \end{aligned}$$

*Indeed, for (14) $\Longrightarrow$ (15) we refer to Remark 2 and for (15) $\Longrightarrow$ (16) to Remark 3. That (16) implies (14) can be seen as follows: For arbitrary $\xi \in \mathbb{R}^N$, choose $h(x) = \xi \cdot x$ in (10). Using (13), (10) turns into $\xi \cdot A^{-1}\xi \leq \frac{1}{\rho}|\xi|^2$, which amounts to (14).*

*We now give the argument for optimality: Let $\mu$ satisfy LSI$(\rho)$. By (15) $\Longrightarrow$ (14) we must have*

$$A \ \geq \ \rho\, \mathrm{id} \quad \textit{in the sense of quadratic forms.}$$

*Because of (14) $\Longrightarrow$ (15), we have for every* $i \in \{1, \cdots, N\}$*:*

$$\begin{aligned}
&\forall\, (x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_N) \\
&\mu(dx_i | x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_N) \quad \textit{satisfies LSI}(A_{ii}).
\end{aligned}$$

*Finally, in view of (12), we have for every* $i < j \in \{1, \cdots, N\}$*:*

$$\forall\, (x_1, \cdots, x_N) \quad |\nabla_i \nabla_j H(x_1, \cdots, x_N)| \;=\; -A_{ij}.$$

**Remark 5 (Royer's Théorème 5.2.1)** *In this remark, we compare our Theorem 1 to Royer's [R, Théorème 5.2.1]. In view of Lemma 6 below, Royer's hypothesis* $(H_2)$ *on the coupling can be rephrased in our language (2) & (3) as*

$$\forall\, i \in \{1, \cdots, N\} \quad \sum_{j \neq i} \frac{\kappa_{ij}}{\rho_i} \;\leq\; \gamma, \tag{17}$$

*for some constant* $\gamma < 1$*, where* $\kappa_{ij} = \kappa_{ji}$*. His hypothesis on the single–site conditional measures in our language (3) turns into*

$$\forall\, i \in \{1, \cdots, N\} \quad \rho_i \;\geq\; \frac{2}{c_0}, \tag{18}$$

*for some constant* $c_0 > 0$*. His result translates into*

$$\mu(dx_1 \cdots dx_N) \quad \textit{satisfies LSI}\Big((1-\gamma)^2\, \frac{2}{c_0}\Big). \tag{19}$$

*(In fact, there seems to be a typo in the statement of his result [R, (87)]. The inequality (19) is taken from [R, (92)] in Royer's proof, which is stronger than the actual statement [R, (87)] by a factor of two.)*

*On the other hand, (17) and (18) imply that the matrix* $A$ *defined in (4) satisfies in the sense of quadratic forms*

$$A \;\geq\; \frac{2}{c_0}(1-\gamma)\,\mathrm{id}. \tag{20}$$

*Indeed, consider the smallest eigenvalue* $\rho$ *of* $A$*. Let* $x$ *denote a corresponding eigenvector. Let* $i \in \{1, \cdots, N\}$ *be such that* $|x_i| = \max_j |x_j|$*. W. l. o. g. we may assume*

$$x_i \;=\; \max_j |x_j| \;=\; 1. \tag{21}$$

*Then the* $i$*–th component of the identity* $\rho x = Ax$ *reads*

$$\begin{aligned}
\rho &\overset{(21)}{=} \rho_i - \sum_{j\neq i} \kappa_{ij} x_j \\
&\geq \rho_i - \left(\sum_{j\neq i} \kappa_{ij}\right) \max_k |x_k| \\
&\overset{(21)}{=} \rho_i - \sum_{j\neq i} \kappa_{ij} \\
&\overset{(17)}{\geq} \rho_i - \gamma\, \rho_i \\
&\overset{(18)}{\geq} \frac{2}{c_0}(1-\gamma).
\end{aligned}$$

*This establishes (20). Hence Theorem 1 implies that*

$$\mu(dx_1 \cdots dx_N) \quad \textit{satisfies LSI}\left((1-\gamma)\,\frac{2}{c_0}\right),$$

*which is stronger than Royer's by a factor of* $(1-\gamma)$*.*

### *1.2 Two–scale theorem*

We will apply Theorem 1 in conjunction with a "two–scale criterion" for LSI, implicitly contained in [GORV] Proposition 2 and [BB] Theorem 1.3. The two–scale criterion states that LSI for a conditional measure and the corresponding marginal may be combined – regardless of the interaction strength – to prove a global LSI. We state the result in the case of a product space:

**Theorem 2** *Let* $X_1$*,* $X_2$ *be Euclidean spaces and* $\mu(dx_1 dx_2)$ *a probability measure on the product space* $X_1 \times X_2$ *with a smooth, positive Lebesgue density*

$\frac{d\mu}{d\mathcal{L}}$.

*We assume that there exists* $\kappa_{12} < \infty$ *such that the Hamiltonian* $H(x_1, x_2) = -\log \frac{d\mu}{d\mathcal{L}}$ *satisfies*

$$\forall\, (x_1, x_2) \quad |\nabla_1 \nabla_2 H(x_1, x_2)| \;\le\; \kappa_{12}. \tag{22}$$

*We assume that there exist* $\rho_2 > 0$*,* $\bar{\rho}_1 > 0$ *such that we have for the conditional measure and the marginal:*

$$\forall x_1 \in X_1 \quad \mu(dx_2|x_1) \ \textit{satisfies} \ LSI(\rho_2), \tag{23}$$

$$\bar{\mu}(dx_1) \ \textit{satisfies} \ LSI(\bar{\rho}_1). \tag{24}$$

*Then we obtain*

$$\begin{aligned} &\mu(dx_1 dx_2) \ \textit{satisfies} \ LSI(\rho) \ \textit{with} \\ \rho \;\ge\; &\frac{1}{2}\left(\rho_2 + \bar{\rho}_1 + \frac{\kappa_{12}^2}{\rho_2} - \left(\left(\rho_2 + \bar{\rho}_1 + \frac{\kappa_{12}^2}{\bar{\rho}_1}\right)^2 - 4\rho_2\,\bar{\rho}_1\right)^{1/2}\right). \end{aligned} \tag{25}$$

**Remark 6 (Product measures)** *In the absence of coupling, i.e.* $\kappa_{12} = 0$*, we recover from* (25) *the product estimate*

$$\rho \;\ge\; \frac{1}{2}\left(\rho_2 + \bar{\rho}_1 - |\rho_2 - \bar{\rho}_1|\right) \;=\; \min\{\rho_2, \bar{\rho}_1\}.$$

**Remark 7 (Linear algebra)** *The expression on the right–hand side of* (25) *has a linear algebra characterization: For* $\rho_2,\ \bar{\rho}_1 > 0$ *and* $\kappa_{12} < \infty$*, let* $\rho$ *be the maximal constant with the following property. For all symmetric* $2\times 2$ *matrices*

$$\bar{A} \;=\; \begin{pmatrix} A_{11} & A_{12} \\ \\ A_{12} & A_{22} \end{pmatrix}$$

*with*

$$|A_{12}| \;\leq\; \kappa_{12}, \qquad A_{22} \;\geq\; \rho_2, \qquad A \;\geq\; \bar{\rho}_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{26}$$

*we have*

$$A \;\geq\; \rho \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{27}$$

*Then* $\rho$ *is given by the expression on the right–hand side of* (25)*.*

*Indeed, the last property in* (26) *is equivalent to*

$$A_{11} - \frac{A_{12}^2}{A_{22}} \;\geq\; \bar{\rho}_1. \tag{28}$$

*We notice that the smallest eigenvalue of* $A$ *is given by*

$$\rho_{min} \;=\; \frac{1}{2}\left(A_{11} + A_{22} - \sqrt{(A_{11} - A_{22})^2 + 4A_{12}^2}\right). \tag{29}$$

*Since this expression is monotone increasing in* $A_{11}$*, we have by* (28) *that*

$$\rho_{min} \;\geq\; \frac{1}{2}\left(\bar{\rho}_1 + \frac{A_{12}^2}{A_{22}} + A_{22} - \sqrt{\left(\bar{\rho}_1 + \frac{A_{12}^2}{A_{22}} - A_{22}\right)^2 + 4A_{12}^2}\right),$$

*with equality if there is equality in* (28)*. Since the expression*

$$\frac{1}{2}\left(\bar{\rho}_1 + x + y - \sqrt{(\bar{\rho}_1 + x - y)^2 + 4xy}\right)$$

*is monotone decreasing in* $x = A_{12}^2/A_{22}$ *and monotone increasing in* $y = A_{22}$*, we have in view of* (26)

$$\rho_{min} \;\geq\; \frac{1}{2}\left(\bar{\rho}_1 + \frac{\kappa_{12}^2}{\rho_2} + \rho_2 - \sqrt{\left(\bar{\rho}_1 + \frac{\kappa_{12}^2}{\rho_2} - \rho_2\right)^2 + 4\kappa_{12}^2}\right),$$

*with equality if there is equality in* (26) *and* (28)*.*

**Remark 8 (Gaussians)** *Assume* $X_1 = X_2 = \mathbb{R}$ *and let* $A$ *be a symmetric, positive definite* $2 \times 2$ *matrix. In this remark, we argue that Theorem 2 is optimal for Gaussians, i.e. for Hamiltonians of the form*

$$H \;=\; -\log\frac{d\mu}{d\mathcal{L}} \;=\; \frac{1}{2}\begin{pmatrix} x_1 \\ \\ x_2 \end{pmatrix}\cdot\begin{pmatrix} A_{11} & A_{12} \\ \\ A_{12} & A_{22} \end{pmatrix}\begin{pmatrix} x_1 \\ \\ x_2 \end{pmatrix} - \begin{pmatrix} b_1 \\ \\ b_2 \end{pmatrix}\cdot\begin{pmatrix} x_1 \\ \\ x_2 \end{pmatrix}.$$

*Indeed, the assumption* (22) *turns into*

$$|A_{12}| \;\le\; \kappa_{12},$$

*and in view of the equivalence* (14) $\Longleftrightarrow$ (15)*, the assumption* (23) *translates into*

$$A_{22} \;\ge\; \rho_2.$$

*Similarly, since* $\bar{\mu}(dx_1)$ *is a Gaussian with Hamiltonian*

$$\bar{H} \;=\; -\log\frac{d\bar{\mu}}{d\mathcal{L}} \;=\; \frac{1}{2}x_1\left(A_{11} - \frac{A_{12}^2}{A_{22}}\right)x_1 + \bar{b}_1 x_1,$$

*assumption* (24) *translates into*

$$A_{11} - \frac{A_{12}^2}{A_{22}} \;\ge\; \bar{\rho}_1.$$

We now appeal to Remark 7 (including (28)).

## 2 Two applications

### *2.1 Application I: Weak interactions*

Our first example is a straightforward application of Theorem 1 in the case of non–convex potential and weak interactions. This problem is quite well–known in the literature and has been analyzed both by the methods of Zegarlinski, for instance in [R,BH], and by the method of Lu–Yau, for instance in [Yo1]. A nice treatment is given in [L, Theorem 6.3]. Studying the problem of weak interactions via Theorem 1 leads to a particularly simple analysis; it is to illustrate this point that we include the application.

We consider nearest–neighbor interactions in two dimensions; without difficulty one can generalize to higher dimensions and either finite–range interactions or infinite–range interactions that decay sufficiently quickly.

Let $X$ denote a periodic $N$–site lattice in two dimensions. We assume that $\mu \in \mathcal{P}(X)$ has Hamiltonian

$$H(x) \;=\; \sum_i \psi(x_i) - \varepsilon \sum_{i\sim j} x_i\, x_j, \tag{30}$$

where $i \sim j$ represent nearest neighbor sites and the smooth potential $\psi$ is a bounded perturbation of a Gaussian in the sense that

$$\psi(x) \;=\; \frac{1}{2}\,x^2 + \delta\psi(x), \qquad \text{with} \qquad \sup_{\mathbb{R}} |\delta\psi(x)| < \infty. \tag{31}$$

Define $\Delta \;:=\; \exp(-\mathrm{osc}_{\mathbb{R}}\,\delta\psi)$.

If the interaction is sufficiently weak in the sense that

$$\varepsilon \;\le\; \frac{\Delta}{4}, \tag{32}$$

then

$$\mu \qquad \text{satisfies} \qquad \mathrm{LSI}(\Delta - 4\varepsilon).$$

Let $X_i$ denote a single site on the lattice, so that $X = X_1 \times \cdots \times X_N$. Then the single–site conditional measure

$$\mu_i(dx_i) \;:=\; \mu(dx_i | x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_N) \quad \text{satisfies LSI}(\Delta).$$

To see this, notice that the conditional measure is

$$\mu_i(dx_i) \;=\; \mathcal{Z}^{-1} \exp\Big( -\psi(x_i) + \varepsilon x_i \sum_{j;j\sim i} x_j \Big)\, dx_i.$$

Here $j; j \sim i$ is the set of sites $j$ that are nearest neighbors of $i$. Thus, the measure is a perturbed Gaussian:

$$\mu_i(dx_i) \;=\; \mathcal{Z}^{-1} \exp\Big( -\frac{1}{2} x_i^2 + \varepsilon x_i \sum_{j;j\sim i} x_j - \delta\psi(x_i) \Big)\, dx_i.$$

Since by the Bakry–Emery Principle the Gaussian satisfies LSI(1) (cf. [BE] and also Remark 2, above), we have by the Holley–Stroock principle [HS] that

$$\mu_i(dx_i) \quad \text{satisfies} \quad \mathrm{LSI}\Big( \exp(-\mathrm{osc}_{\mathbb{R}}\, \delta\psi) \Big). \tag{33}$$

Finally, observe that $|\nabla_i \nabla_j H| = \varepsilon$ for $i$ and $j$ nearest neighbors (there are four in two dimensions), and zero otherwise.

Thus, we may invoke Theorem 1. Given (32), the matrix $A$ satisfies

$$A \;\ge\; (\Delta - 4\varepsilon)\, \mathrm{id}.$$

### *2.2 Application II: Strong interactions*

In our second application, we again treat a spin system with continuous, unbounded spins on a multi–dimensional lattice. This time we consider the non–perturbative case. This means that instead of assuming weak interactions, we assume the exponential decay of correlations of the spin variables in the distance between the corresponding sites, cf. (37). In equilibrium statistical mechanics, such a property is associated with the absence of phase transition. In this sense, Theorem 3 relates an equilibrium property (spatial decay of correlations) to a dynamical property (relaxation time to equilibrium for Glauber dynamics, as estimated by the constant in LSI). The first results of this type were obtained by Stroock and Zegarlinski in the case of a bounded continuous spin space [SZ] and (by a different method) in the case of a discrete spin space [SZ2].

The case of an unbounded continuous spin space was first treated in [Z3]. However while [Z3] treats the case of a one–dimensional lattice comprehensively, the result for a multi–dimensional lattice is just stated ([Z3] Theorem 5.1), referring the reader to [SZ2]. A full proof was given by Yoshida in [Yo1] under a less decay of correlations conditions ([Yo1, (2.1) in Theorem 2.1]). Thanks to an idea by Bodineau and Hellfer [BH], a more transparent criterion is given in [Yo2, (DS1) in Theorem 2.1].

Let us comment a bit on the decay of correlation condition (37). Notice that (37) is required for all subsets $\Lambda \setminus S$ of sites uniformly in the "boundary conditions" $(x_k)_{k \in S}$. Conditions of this type have been scrutinized by Martinelli and Olivieri ([MO1, Section 2]): They do not cover the entire one–phase region. In

[MO2], a less demanding condition is shown to be sufficient in the case of a discrete spin space.

Finally, we remark that we assume superquadratic growth of the single–site potential, cf. (34), as in [Yo1,Yo2]. We credit the present version of the result to Yoshida, but offer an independent proof based on our Theorems 1 and 2.

**Theorem 3** *Let $\Lambda \subset \mathbb{Z}^d$ be an arbitrary, finite subset. Let $\psi$ be a function of the form*

$$\psi(x) \;=\; \frac{1}{12}x^4 + \delta\psi(x) \qquad \text{with} \quad |\delta\psi''(x)| \;\leq\; C. \tag{34}$$

*Let $(J_{ij})_{i,j\in\Lambda}$ be a symmetric matrix with $J_{ii} = 0$. Consider the Hamiltonian on $\mathbb{R}^\Lambda$*

$$H((x_i)_{i\in\Lambda}) \;=\; \sum_{i\in\Lambda}\psi(x_i) - \frac{1}{2}\sum_{i\in\Lambda}\sum_{j\in\Lambda} J_{ij}x_ix_j \tag{35}$$

*and the corresponding Gibbs measure*

$$d\mu \;=\; \frac{1}{Z}\exp(-H)\;\Pi_{i\in\Lambda}dx_i.$$

*We assume the uniform control:*

$$|J_{ij}| \;\leq\; C\exp(-|i-j|/C) \tag{36}$$

*for $i, j \in \Lambda$ and*

$$|\langle x_i; x_j\rangle_S| \;\leq\; C\exp(-|i-j|/C) \tag{37}$$

*for $S \subset \Lambda$ and $i, j \in \Lambda\setminus S$, where $\langle\cdot;\cdot\rangle_S$ denotes the covariance with respect to the conditional measure*

$$\mu\left(\Pi_{i\in\Lambda\setminus S}dx_i \Big| (x_k)_{k\in S}\right).$$

*Then*

$$\mu \qquad \text{satisfies} \qquad \mathrm{LSI}(\rho)$$

*with a constant* $\rho > 0$ *depending only on* $C$ *and* $d$*.*

The proof of Theorem 3 will follow from a series of four lemmas. The first lemma relies on the superquadratic growth (34) of the single–spin potential and follows from a combination of the criteria of Bakry & Emery and Holley & Stroock. A proof was given in [Yo1, Lemma 3.2]. For the convenience of the reader, we reproduce it here.

**Lemma 1** *For every* $C < \infty$ *there exists* $\tilde{C} < \infty$ *such that if* $\psi$ *satisfies* (34)*, then*

$$\mu(dx) \;:=\; \frac{1}{Z}\exp\left(-\psi(x)\right)\,dx \qquad \text{satisfies} \qquad \mathrm{LSI}(1/\tilde{C}).$$

The second lemma converts the information on the covariance into information on the coarse–grained Hamiltonian.

**Lemma 2 (Coarse–grained Hamiltonian)** *There exists* $\bar{C} < \infty$ *depending only on* $d$ *and the constant* $C$ *in* (36)*,* (37) *with the following property: Let* $S$ *be an arbitrary non–empty subset of the set* $\Lambda \subset \mathbb{Z}^d$ *of sites. Consider the related marginal* $\bar{\mu}\left(\Pi_{i\in S}\,dx_i\right)$ *and its Hamiltonian* $\bar{H}((x_i)_{i\in S})$*. Then we have*

$$\left|\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}\bar{H} - \psi''(x_i)\delta_{ij}\right| \;\le\; \bar{C}\exp(-|i-j|/\bar{C}) \tag{38}$$

*for all* $i, j \in S$*.*

This is enough to deduce LSI for any single–site conditional measure:

**Lemma 3 (Single–site conditionals)** *There exists* $\rho > 0$ *depending only on* $d$ *and the constant* $C$ *in* (34)*,* (36)*, and* (37) *with the following property: Let the subset* $S \subset \Lambda$ *and the site* $i \in S$ *be arbitrary. Consider the distribution*

$\bar{\mu}(dx_i|(x_j)_{j\in S\setminus\{i\}})$ *of* $x_i$ *conditioned on* $(x_j)_{j\in S\setminus\{i\}}$. *Then*

$$\bar{\mu}(dx_i|(x_j)_{j\in S\setminus\{i\}}) \qquad \text{satisfies} \qquad \mathrm{LSI}(\rho)$$

*uniformly in* $(x_j)_{j\in S\setminus\{i\}}$.

Finally, Theorem 1 together with Lemmas 2 and 3 allows us to prove LSI for measures on the $K$–sublattice $\Lambda_K$ (cf. (39) below and Figure 1).

**Lemma 4 ($K$–sublattice)** *There exists a constant* $\bar{\rho} > 0$ *and a* $K \in \mathbb{N}$ *depending only on* $d$ *and the constants* $C$ *in* (34), (36), *and* (37) *with the following property: Let* $i_0 \in \mathbb{Z}^d$ *be arbitrary and consider the* $K$*–sublattice:*

$$\Lambda_K \ := \ (i_0 + K\mathbb{Z}^d) \cap \Lambda. \tag{39}$$

*Let* $S$ *be an arbitrary set with* $\Lambda_K \subseteq S \subseteq \Lambda$. *Consider the distribution* $\bar{\mu}(\Pi_{i\in\Lambda_K} dx_i|(x_j)_{j\in S\setminus\Lambda_K})$ *of* $(x_i)_{i\in\Lambda_K}$ *conditioned on* $(x_j)_{j\in S\setminus\Lambda_K}$. *Then*

$$\bar{\mu}\left(\Pi_{i\in\Lambda_K} dx_i|(x_j)_{j\in S\setminus\Lambda_K}\right) \qquad \text{satisfies} \qquad \mathrm{LSI}(\bar{\rho})$$

*uniformly in* $(x_j)_{j\in S\setminus\Lambda_K}$.

PROOF OF THEOREM 3.

To prove LSI for the full measure $\mu$, we now apply the two scale criterion from Theorem 2 to a sequence of conditional and marginal measures (cf. Figures 1 through 4).

Before explaining the sequence, it is convenient to introduce some notation. Fix $i_0$ and let $\Lambda_K^\ell$ for $\ell = 0, \ldots, K^d - 1$ be an enumeration of the translates of $\Lambda_K$:

$$(\Lambda_K + c_1\,\mathbf{e}_1 + c_2\mathbf{e}_2 + \ldots + c_d\mathbf{e}_d) \cap \Lambda, \qquad \text{for } c_1, c_2, \ldots c_d \in 0, \ldots, K-1,$$

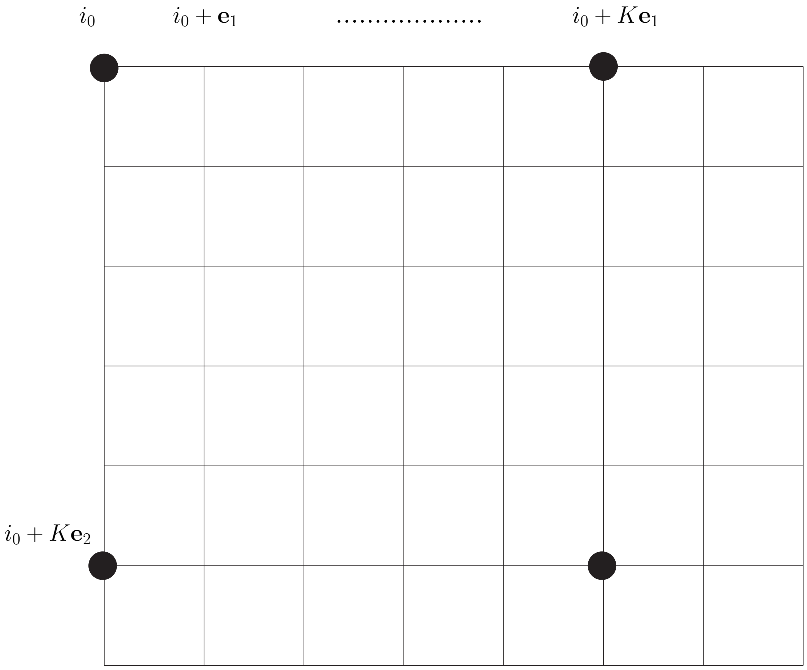

Fig. 1. Illustration of the $K$–sublattice for $d = 2$. We establish LSI for $\mu_c^{(1)}$, the measure that is "active" on every $K^{th}$ site (black circles) and conditioned on the spin at every other site (no circles drawn).

where $\mathbf{e}_1, \ldots, \mathbf{e}_d$ are unit vectors that generate $\mathbb{Z}^d$. Let

$$L_K^n \;=\; \cup_{\ell=0}^n \Lambda_K^\ell .$$

By definition, $\Lambda_K^0 = L_K^0 = \Lambda_K$.

Finally, define

$$\mu_c^{(n)} \;:=\; \mu\left(\Pi_i dx_i, i \in L_K^{n-1} \mid (x_j)_{j \notin L_K^{n-1}}\right)$$

and

$$\bar{\mu}^{(n)} \;:=\; \bar{\mu}\left(\Pi_i dx_i, i \in \Lambda_K^n \mid (x_j)_{j \notin L_K^n}\right).$$

By Lemma 4 with $\Lambda_K = \Lambda_K^0$ and $S = \Lambda$, we have that $\mu_c^{(1)}$ satisfies LSI($\bar{\rho}$). By Lemma 4 with $\Lambda_K = \Lambda_K^1$ and $S = \Lambda \setminus L_K^0$, $\bar{\mu}^{(1)}$ satisfies LSI($\bar{\rho}$). By Theorem 2, these two ingredients imply LSI for $\mu_c^{(2)}$. (See Figures 1, 2, and 3 for

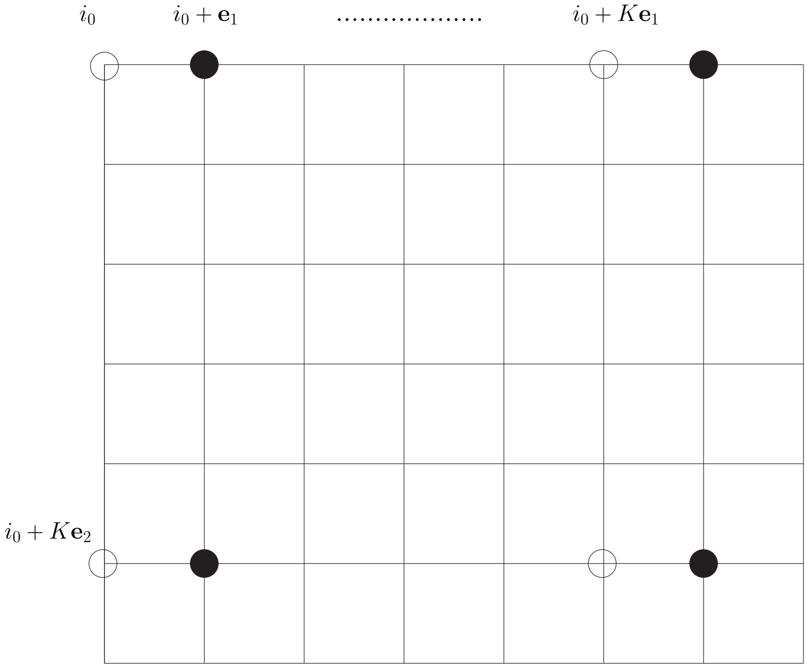

Fig. 2. We establish LSI for $\bar{\mu}^{(1)}$, the measure that is "active" on the black circles, averaged over the open circles, and conditioned on the spin at the other sites.

an illustration.)

We now use this information, Lemma 4 for $\bar{\mu}^{(2)}$, and Theorem 2 to conclude LSI for $\mu_c^{(3)}$. We continue in this way until having proved LSI for

$$\mu_c^{(K^d)} = \mu.$$

### *2.2.1 Proofs of the lemmas*

Proof of Lemma 1.

The main idea is to show that $\psi$ can be written as

$$\psi(x) \; = \; \psi_c(x) + \psi_b(x) \tag{40}$$

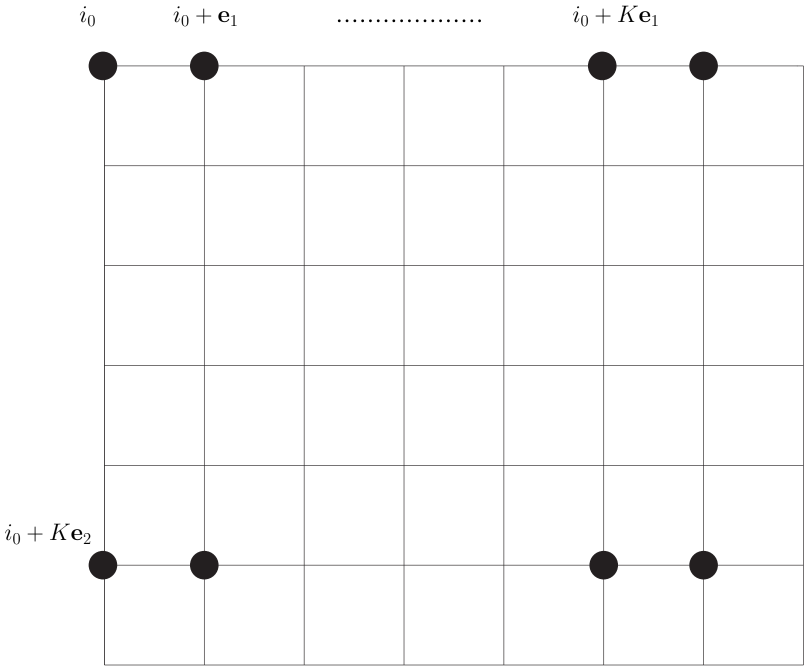

Fig. 3. We use the two–scale criterion from Theorem 2 to deduce from LSI for $\mu_c^{(1)}$ and $\bar{\mu}^{(1)}$ the LSI for $\mu_c^{(2)}$, the measure that is “active” on the black circles.

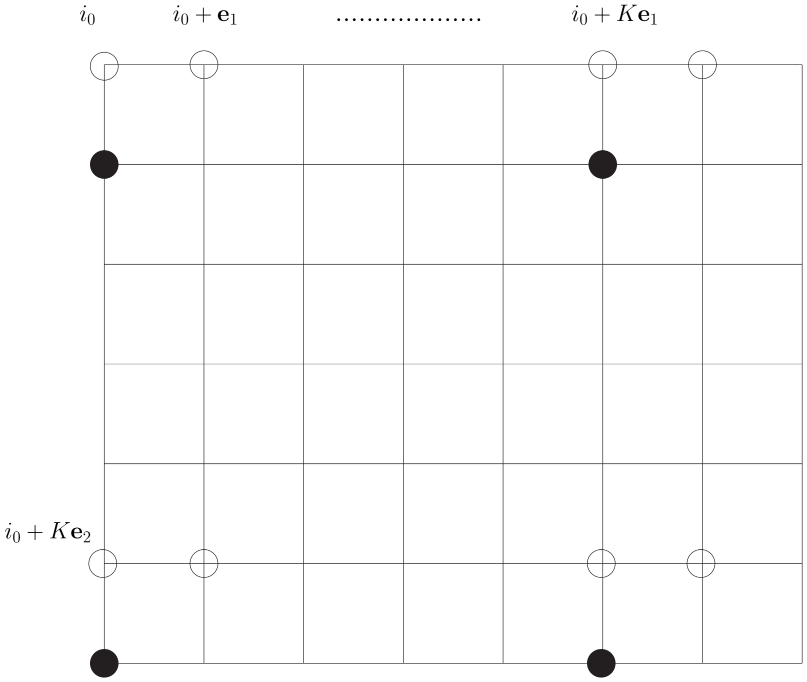

Fig. 4. This illustration depicts $\bar{\mu}^{(2)}$.

with

$$\psi_c''(x) \ \geq \ 1 \tag{41}$$

and

$$|\psi_b(x)| \ \leq \ C_1, \tag{42}$$

with $C_1$ depending only on $C$ in (34). Both properties (34) and (41) are invariant under the modification of $\psi$ by an affine function. Hence, we may assume without loss of generality that $\psi(0) = \psi'(0) = 0$, which in view of (34) translates into

$$\delta\psi(0) \ = \ \delta\psi'(0) \ = \ 0. \tag{43}$$

We make the following ansatz:

$$\begin{aligned} \psi_c(x) \ &= \ \psi(x) + \eta_R(x)\left(\frac{1}{2}x^2 - \delta\psi(x)\right) \\ &\overset{(34)}{=} \ \frac{1}{12}x^4 + \eta_R(x)\frac{1}{2}x^2 + (1-\eta_R(x))\delta\psi(x), \end{aligned} \tag{44}$$

and thus, to satisfy (40), we must have

$$\psi_b(x) \ = \ \psi(x) - \psi_c(x) \ = \ \eta_R(x)\left(\delta\psi(x) - \frac{1}{2}x^2\right). \tag{45}$$

Here $\eta_R$ is a cut–off function of the form $\eta_R(x) = \eta(x/R)$, where $\eta \in [0,1]$ is such that

$$\eta(x) \ = \ 1 \ \text{ for } \ |x| \ \leq \ 1 \qquad \text{and} \qquad \eta(x) \ = \ 0 \ \text{ for } \ |x| \ \geq \ 2. \tag{46}$$

Let $M < \infty$ denote a generic universal constant. We have by (44):

$$\psi_c'' \ = \ x^2 + \eta_R + (1-\eta_R)\delta\psi'' + 2\eta_R'(x - \delta\psi') + \eta_R''\Big(\frac{1}{2}x^2 - \delta\psi\Big),$$

so that in view of (46) we obtain for $|x| \geq 2R$:

$$\psi_c'' \;=\; x^2 + \delta\psi'' \;\geq\; 4R^2 - C, \tag{47}$$

for $|x| \leq R$:

$$\psi_c'' \;=\; x^2 + 1 \;\geq\; 1, \tag{48}$$

and for $R \leq |x| \leq 2R$:

$$\begin{aligned}\psi_c'' \;&\geq\; R^2 - \sup_{x\in\mathbb{R}} |\delta\psi''| - \frac{M}{R}(2R + \sup_{|x|\leq 2R} |\delta\psi'|) - \frac{M}{R^2}(2R^2 + \sup_{|x|\leq 2R} |\delta\psi|) \\ &\overset{(43)}{\geq}\; R^2 - M(1+C). \end{aligned} \tag{49}$$

From (47), (48), and (49), it follows that for

$$R^2 \geq \max\{(1+C)/4, 1 + M(1+C)\},$$

we have as desired

$$\psi_c'' \;\geq\; 1.$$

We now turn to $\psi_b$, cf. (45). For $|x| \geq 2R$, we have $\psi_b(x) = 0$, whereas for $|x| \leq 2R$,

$$|\psi_b(x)| \;\leq\; \sup_{|x|\leq 2R} |\delta\psi| + 2R^2 \;\leq\; C + 2R^2,$$

so that

$$\sup_{x\in\mathbb{R}} |\psi_b(x)| \;\leq\; C + 2R^2 \;=\; C + M =: C_1.$$

We now apply the criteria of Bakry–Emery ([BE], and see also Remark 2 above) and Holley–Stroock [HS]. By Bakry–Emery, (41) implies that

$$\tilde{\mu}(dx) \;=\; \frac{1}{\tilde{Z}}\exp(-\psi_c(x))dx \qquad \text{satisfies} \qquad \mathrm{LSI}(1).$$

By Holley–Stroock, (42) implies that

$$\mu(dx) \ = \ \exp(-\psi_c(x) - \psi_b(x))\, dx \qquad \text{satisfies} \qquad \mathrm{LSI}(1/\tilde{C})$$

with

$$\tilde{C} \ = \ \exp(\mathrm{osc}\psi_b)) \ \leq \ \exp(2C_1).$$

Proof of Lemma 2.

The starting point is the relation between the original Hamiltonian $H$ and the coarse–grained Hamiltonian

$$\bar{H}((x_i)_{i\in S}) \ = \ -\log\int \exp(-H(x))\Pi_{k\in\Lambda\setminus S} dx_k,$$

which we differentiate to obtain:

$$\frac{\partial}{\partial x_j}\bar{H} \ = \ \frac{\displaystyle\int \frac{\partial}{\partial x_j} H \exp(-H)\ \Pi_{k\in\Lambda\setminus S}\ dx_k}{\displaystyle\int \exp(-H)\ \Pi_{k\in\Lambda\setminus S}\ dx_k},$$

and:

$$\begin{aligned}
&\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}\bar{H} \\
&= \frac{\displaystyle\int \frac{\partial}{\partial x_j}\frac{\partial}{\partial x_j} H \exp(-H)\ \Pi_{k\in\Lambda\setminus S}\ dx_k}{\displaystyle\int \exp(-H)\ \Pi_{k\in\Lambda\setminus S}\ dx_k} \\
&- \frac{\displaystyle\int \frac{\partial}{\partial x_i} H \frac{\partial}{\partial x_j} H \exp(-H)\ \Pi_{k\in\Lambda\setminus S}\ dx_k}{\int \exp(-H)\Pi_{k\in\Lambda\setminus S}\ dx_k} \\
&+ \frac{\displaystyle\int \frac{\partial}{\partial x_i} H \exp(-H)\Pi_{k\in\Lambda\setminus S} dx_k \int \frac{\partial}{\partial x_j} H \exp(-H)\Pi_{k\in\Lambda\setminus S}\ dx_k}{\left(\displaystyle\int \exp(-H)\ \Pi_{k\in\Lambda\setminus S}\ dx_k\right)^2}.
\end{aligned} \tag{50}$$

If $\langle\cdot\rangle$ and $\langle\cdot;\cdot\rangle$ denote expectation and covariance with respect to the conditional measure $\mu(\Pi_{k\in\Lambda\setminus S}\, dx_i\big|(x_i)_{i\in S})$, then (50) can be formulated as

$$\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}\bar{H} \;=\; \left\langle \frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}H\right\rangle - \left\langle \frac{\partial}{\partial x_i}H\,;\,\frac{\partial}{\partial x_j}H\right\rangle. \tag{51}$$

For our Hamiltonian (35), we have

$$\begin{aligned}\frac{\partial}{\partial x_i}H &= \psi'(x_i) - \sum_{k\in\Lambda} J_{ik}x_k,\\ \frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}H &= \psi''(x_i)\delta_{ij} - J_{ij}.\end{aligned}$$

Since we condition on $(x_i)_{i\in S}$, we have

$$\begin{aligned}\langle\psi''(x_i)\rangle &= \psi''(x_i),\\ \left\langle \psi'(x_i) - \sum_{k\in\Lambda}J_{ik}x_k\,;\,\psi'(x_j) - \sum_{\ell\in\Lambda}J_{j\ell}x_\ell\right\rangle &= \sum_{k\in\Lambda\setminus S}\ \sum_{\ell\in\Lambda\setminus S} J_{ik}J_{j\ell}\langle x_k\,;\,x_\ell\rangle,\end{aligned}$$

so that (51) can be reexpressed as

$$\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}\bar{H} - \psi''(x_i)\delta_{ij} \;=\; -J_{ij} - \sum_{k\in\Lambda\setminus S}\ \sum_{\ell\in\Lambda\setminus S} J_{ik}J_{j\ell}\langle x_k\,;\,x_\ell\rangle.$$

This identity is the basis for our estimate:

$$\left|\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}\bar{H} - \psi''(x_i)\delta_{ij}\right| \;\le\; |J_{ij}| + \sum_{k\in\Lambda\setminus S}\ \sum_{\ell\in\Lambda\setminus S} |J_{ik}||J_{j\ell}||\langle x_k\,;\,x_\ell\rangle|. \tag{52}$$

We now appeal to the triangle inequality

$$|i-j| \;\le\; |i-k| + |j-\ell| + |k-\ell|, \tag{53}$$

which we write as

$$\begin{aligned}\frac{1}{2}(|i-j| + |i-k| + |j-\ell|) &\overset{(53)}{\le} |i-k| + |j-\ell| + \frac{1}{2}|k-\ell|\\ &\le |i-k| + |j-\ell| + |k-\ell|,\end{aligned}$$

to see that

$$\begin{aligned}
&\sum_{k\in\mathbb{Z}^d}\sum_{\ell\in\mathbb{Z}^d}\exp\left(-(|i-k|+|j-\ell|+|k-\ell|)/C\right)\\
&\le\ \exp(-|i-j|/2C)\sum_{k\in\mathbb{Z}^d}\exp(-|i-k|/2C)\sum_{\ell\in\mathbb{Z}^d}\exp(-|j-\ell|/2C)\\
&\le\ \exp(-|i-j|/2C)\left(\sum_{k\in\mathbb{Z}^d}\exp(-|k|/2C)\right)^2.
\end{aligned}$$

Hence (52) together with (36) and (37) yields

$$\begin{aligned}
&\left|\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}\bar{H}-\psi''(x_i)\delta_{ij}\right|\\
&\le\ C\left(1+\left(C\sum_{k\in\mathbb{Z}^d}\exp(-|k|/2C)\right)^2\right)\exp(-|i-j|/2C),
\end{aligned}$$

so that (38) holds with

$$\bar{C}=C\max\left\{1+\left(C\sum_{k\in\mathbb{Z}^d}\exp(-|k|/2C)\right)^2,2\right\}.$$

PROOF OF LEMMA 3.

Let $\bar{H}$ be as in Lemma 2. Fix $(x_j)_{j\in S\setminus\{i\}}$, consider $\bar{H}$ as a function of $x_i$, and define the self–potential:

$$\bar{\psi}(x_i)\ :=\ \bar{H}(x_i,(x_j)_{j\in S\setminus\{i\}}).$$

Then $\bar{\mu}(dx_i\,|\,(x_j)_{j\in S\setminus\{i\}})$ is of the form

$$\bar{\mu}(dx_i\,|\,(x_j)_{j\in S\setminus\{i\}})\ =\ \frac{1}{Z}\exp(-\bar{\psi}(x_i))\,dx_i.$$

We learn from Lemma 2 that

$$|\bar{\psi}''(x_i)-\psi''(x_i)|\ \le\ \bar{C},$$

which means that according to (34),

$$\bar{\psi}(x_i) \;=\; \frac{1}{12}x^4 + \delta\bar{\psi}(x_i),$$

with

$$|\delta\bar{\psi}''(x_i)| \;\le\; C + \bar{C}.$$

Hence by Lemma 1, there exists $\rho > 0$ depending only on $C$ in (34), (36), and (37) such that

$$\bar{\mu}(dx_i \,|\, (x_j)_{j\in S\setminus\{i\}}) \qquad \text{satisfies} \qquad \mathrm{LSI}(\rho).$$

Proof of Lemma 4.

Fix $(x_j)_{j\in S\setminus\Lambda_K}$. We will apply Theorem 1 to the measure

$$\tilde{\mu}\left(\Pi_{i\in\Lambda_K}\, dx_i\right) \;:=\; \bar{\mu}\left(\Pi_{i\in\Lambda_K}\, dx_i \mid (x_j)_{j\in S\setminus\Lambda_K}\right)$$

with its Hamiltonian

$$\tilde{H}\left((x_i)_{i\in\Lambda_K}\right) \;:=\; \bar{H}((x_i)_{i\in S}),$$

where $\bar{H}$ is as in Lemma 2.

For the single–site marginals, we have according to Lemma 3:

$$\tilde{\mu}\left(dx_i \mid (x_j)_{j\in\Lambda_K\setminus\{i\}}\right) \qquad \text{satisfies} \qquad \mathrm{LSI}(\rho)$$

for all $i \in \Lambda_K$.

For the mixed derivatives of the Hamiltonian, we have according to Lemma 2:

$$\left|\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}\tilde{H}\right| \;\le\; \bar{C}\exp(-|i-j|/\bar{C})$$

for all $i \neq j \in \Lambda_K$.

To apply Theorem 1, we have to consider the symmetric matrix $A = (A_{ij})_{i,j \in \Lambda_K}$ with

$$\begin{aligned} A_{ii} &= \rho \\ A_{ij} &= -\bar{C} \exp(-|i-j|/\bar{C}) \qquad \text{for } \ i \neq j, \end{aligned}$$

and argue that it is positive definite.

We notice

$$\begin{aligned} &\sum_{i \in \Lambda_K} \sum_{j \in \Lambda_K} x_i A_{ij} x_j \\ &= \rho \sum_{i \in \Lambda_K} x_i^2 + \sum_{i \in \Lambda_K} x_i \sum_{j \in \Lambda_K \setminus \{i\}} A_{ij} x_j, \end{aligned}$$

and

$$\sum_{i \in \Lambda_k} x_i \sum_{j \in \Lambda_K \setminus \{i\}} A_{ij} x_j \leq \left( \sum_{i \in \Lambda_K} x_i^2 \sum_{i \in \Lambda_K} \Big( \sum_{j \in \Lambda_K \setminus \{i\}} A_{ij} x_j \Big)^2 \right)^{1/2},$$

with

$$\begin{aligned} \sum_{i \in \Lambda_K} \Big( \sum_{j \in \Lambda_K \setminus \{i\}} A_{ij} x_j \Big)^2 &\leq \sum_{i \in \Lambda_K} \left( \sum_{j \in \Lambda_K \setminus \{i\}} |A_{ij}| \sum_{j \in \Lambda_K \setminus \{i\}} |A_{ij}| x_j^2 \right) \\ &\leq \left( \sup_{i \in \Lambda_K} \sum_{j \in \Lambda_K \setminus \{i\}} |A_{ij}| \right) \left( \sum_{j' \in \Lambda_K} x_{j'}^2 \sum_{i' \in \Lambda_K \setminus \{i\}} |A_{i'j'}| \right) \\ &\leq \left( \sup_{i \in \Lambda_K} \sum_{j \in \Lambda_K \setminus \{i\}} |A_{ij}| \right)^2 \sum_{j \in \Lambda_K} x_j^2. \end{aligned}$$

Hence, we have $A \geq \frac{\rho}{2}\,\mathrm{id}$ in the sense of quadratic forms provided that

$$\left( \sup_{i \in \Lambda_K} \sum_{j \in \Lambda_K \setminus \{i\}} |A_{ij}| \right)^2 \leq \frac{\rho}{2}.$$

This is indeed the case for $K > 1$ sufficiently large, since by (36),

$$\sum_{j\in\Lambda_K\setminus\{i\}} |A_{ij}| \;\le\; C \sum_{\ell\in\mathbb{Z}^d\setminus\{0\}} \exp(-K|\ell|/C),$$

which because of

$$K|\ell| \;\ge\; \frac{K}{2}+\frac{K}{2}|\ell| \;\ge\; \frac{K}{2}+|\ell|$$

implies

$$\sum_{j\in\Lambda_K\setminus\{i\}} |A_{ij}| \;\le\; C\exp(-K/2C)\sum_{\ell\in\mathbb{Z}} \exp(-|\ell|/C).$$

## 3 Auxiliary results for Theorem 1

At the core of Theorem 1 is a covariance estimate stated in Lemma 5 below. It goes back to Bodineau & Helffer. Ledoux gave a very efficient proof in [L, Proposition 2.2]. We give yet a different proof which mimics the proof of Talagrand's inequality given in [OV].

**Lemma 5** *Let $\mu(dx)$ be a probability measure on the Euclidean space $X$. We assume that there exists $\rho > 0$ such that*

$$\mu \quad \textit{satisfies LSI}(\rho). \tag{54}$$

*Then we have for arbitrary $f(x) \ge 0$ and $g(x)$:*

$$\begin{aligned}
&\left|\int g\, f\, d\mu - \int g\, d\mu \int f\, d\mu\right| \\
&\le \sup_x |\nabla g| \left(\frac{2}{\rho}\int f\, d\mu \left(\int \Phi(f)\, d\mu - \Phi\left(\int f\, d\mu\right)\right)\right)^{1/2} \\
&\le \sup_x |\nabla g| \,\frac{1}{\rho}\left(\int f\, d\mu \int \frac{1}{f}|\nabla f|^2\, d\mu\right)^{1/2}.
\end{aligned} \tag{55}$$

We also need a linearized version of Lemma 5:

**Corollary 1** *Let $\mu(dx)$ be a probability measure on the Euclidean space $X$. We assume that there exists $\rho > 0$ such that*

$$\mu \quad \textit{satisfies LSI}(\rho).$$

*Then we have for arbitrary $g(x)$ and $h(x)$:*

$$\left| \int g\, h\, d\mu - \int g\, d\mu \int h\, d\mu \right| \;\le\; \frac{1}{\rho} \sup_x |\nabla g| \sup_x |\nabla h|. \tag{56}$$

Lemma 5 will be used to establish the following version of [R, Hypothesis H2, p.91].

**Lemma 6** *Let $X_1, X_2$ be two Euclidean spaces and $\mu(dx_1 dx_2)$ a probability measure on the product space $X_1 \times X_2$ with a smooth positive Lebesgue density $\frac{d\mu}{d\mathcal{L}}$.*

*We assume that there exists $\kappa_{12} < \infty$ such that the Hamiltonian $H(x_1, x_2) = -\log \frac{d\mu}{d\mathcal{L}}$ satisfies*

$$\forall\, (x_1, x_2) \quad |\nabla_1 \nabla_2 H(x_1, x_2)| \;\le\; \kappa_{12}.$$

*We assume that there exists $\rho_2 > 0$ such that we have for the conditional measure*

$$\forall x_1 \quad \mu(dx_2|x_1) \quad \textit{satisfies LSI}(\rho_2).$$

*For arbitrary $f(x_1, x_2) \ge 0$, consider*

$$\bar{f}(x_1) \;=\; \int f(x_1, x_2)\, \mu(dx_2|x_1). \tag{57}$$

*Then we obtain for the marginal $\bar{\mu}(dx_1)$*

$$
\begin{aligned}
&\Big(\int \frac{1}{2\bar{f}}|\nabla_1 \bar{f}|^2\,\bar{\mu}(dx_1)\Big)^{1/2} \\
&\le\ \Big(\int \frac{1}{2f}|\nabla_1 f|^2\,d\mu\Big)^{1/2}\ +\ \frac{\kappa_{12}}{\rho_2}\ \Big(\int \frac{1}{2f}|\nabla_2 f|^2\,d\mu\Big)^{1/2}.
\end{aligned}
$$

The proof of the following lemma, which is based on Lemma 6, amounts to the Lu–Yau martingale method [LY] in the case of only two sites.

**Lemma 7** *Let $X_1, X_2$ be two Euclidean spaces and $\mu(dx_1dx_2)$ a probability measure on the product space $X_1 \times X_2$ with a smooth positive Lebesgue density $\frac{d\mu}{d\mathcal{L}}$.*

*We assume that there exists $\kappa_{12} < \infty$ such that the Hamiltonian $H(x_1, x_2) = -\log\frac{d\mu}{d\mathcal{L}}$ satisfies*

$$\forall\ (x_1, x_2) \quad |\nabla_1\nabla_2 H(x_1, x_2)|\ \le\ \kappa_{12}.$$

*We assume that there exist $\rho_2, \bar{\rho}_1 > 0$ such that we have for the conditional measure and the marginal*

$$
\begin{aligned}
\forall x_1 \quad \mu(dx_2|x_1)\ &\textit{satisfies LSI}(\rho_2), &(58)\\
\bar{\mu}(dx_1)\ &\textit{satisfies LSI}(\bar{\rho}_1). &(59)
\end{aligned}
$$

*Then we obtain for the marginal $\bar{\mu}(dx_2)$*

$$\bar{\mu}(dx_2) \quad \textit{satisfies LSI}(\bar{\rho}_2)$$

*with*

$$\frac{1}{\bar{\rho}_2}\ \le\ \frac{1}{\rho_2} + \frac{1}{\bar{\rho}_1}\frac{\kappa_{12}^2}{\rho_2^2}.$$

The following statement is a simple consequence of Lemma 7. Alternatively, it can be obtained by Zegarlinski's iterative argument, which is outlined for instance in [GZ, Section 5.2].

**Corollary 2** *Let $X_1, X_2$ be two Euclidean spaces and $\mu(dx_1 dx_2)$ a probability measure on the product space $X_1 \times X_2$ with a smooth positive Lebesgue density $\frac{d\mu}{d\mathcal{L}}$.*

*We assume that there exists $\kappa_{12} < \infty$ such that the Hamiltonian $H(x_1, x_2) = -\log \frac{d\mu}{d\mathcal{L}}$ satisfies*

$$\forall\, (x_1, x_2) \quad |\nabla_1 \nabla_2 H(x_1, x_2)| \;\le\; \kappa_{12}.$$

*We assume that there exist $\rho_1, \rho_2 > 0$ such that we have for the conditional measures*

$$\begin{array}{ll} \forall x_2 & \mu(dx_1|x_2) \textit{ satisfies LSI}(\rho_1), \\ \forall x_1 & \mu(dx_2|x_1) \textit{ satisfies LSI}(\rho_2). \end{array}$$

*We assume that*

$$\rho_1\, \rho_2 - \kappa_{12}^2 \;>\; 0. \tag{60}$$

*Then we obtain for the marginal $\bar{\mu}(dx_1)$*

$$\bar{\mu}(dx_1) \quad \textit{satisfies LSI}(\bar{\rho}_1)$$

*with*

$$\bar{\rho}_1 \;\ge\; \rho_1 - \frac{\kappa_{12}^2}{\rho_2}.$$

We will also need the following, which is a consequence of Corollary 1.

**Lemma 8** *Let* $X_1, X_2, X_3$ *be Euclidean spaces and* $\mu(dx_1dx_2dx_3)$ *a probability measure on the product space* $X_1 \times X_2 \times X_3$ *with a smooth positive Lebesgue density* $\frac{d\mu}{d\mathcal{L}}$.

*We assume that for* $i < j \in \{1,2,3\}$ *there exists* $\kappa_{ij} < \infty$ *such that the Hamiltonian* $H(x_1,x_2,x_3) = -\log \frac{d\mu}{d\mathcal{L}}$ *satisfies*

$$\forall\ (x_1,x_2,x_3) \quad |\nabla_i\nabla_j H(x_1,x_2,x_3)| \ \leq\ \kappa_{ij}.$$

*We assume that there exists* $\rho_3 > 0$ *such that we have for the conditional measures*

$$\forall(x_1,x_2) \quad \mu(dx_3|x_1,x_2) \quad \textit{satisfies LSI}(\rho_3).$$

*Consider the Hamiltonian* $\bar{H}(x_1,x_2)$ *belonging to the marginal* $\bar{\mu}(dx_1dx_2)$, *i.e.*

$$\bar{H}(x_1,x_2) \ =\ -\log\int\exp(-H(x_1,x_2,x_3))\,dx_3.$$

*It satisfies*

$$\forall\ (x_1,x_2) \quad |\nabla_1\nabla_2\bar{H}(x_1,x_2)| \ \leq\ \bar{\kappa}_{12}.$$

*with*

$$\bar{\kappa}_{12} \ \leq\ \kappa_{12} + \frac{\kappa_{13}\,\kappa_{23}}{\rho_3}. \tag{61}$$

Finally, we need an elementary result from linear algebra, which we reproduce for convenience.

**Lemma 9** *Consider a symmetric and positive definite matrix* $A$ *with*

$$A_{ij} \ \leq\ 0 \quad \textit{for}\, i < j \in \{1,\cdots,N\}.$$

*Then the inverse matrix* $A^{-1}$ *satisfies*

$$(A^{-1})_{ij} \ \geq\ 0 \quad \textit{for}\, i,j \in \{1,\cdots,N\}.$$

## 4 Proofs

### *4.1 Proofs of auxiliary lemmas*

PROOF OF LEMMA 5.

Without loss of generality, we may assume

$$\int f\,d\mu \;=\; 1.$$

The second inequality in (55) follows from the first and (54). In order to prove the first, we introduce the semigroup $P_t$ related to $\mu$ and defined by

$$P_0 f = f, \tag{62}$$

$$\forall\ g(x) \quad \frac{d}{dt}\int g\,P_t f\,d\mu = -\int \nabla g\cdot\nabla P_t f\,d\mu. \tag{63}$$

All we need to know is

$$\int P_t f\,d\mu = \int f\,d\mu \;=\; 1, \tag{64}$$

$$\frac{d}{dt}\int \Phi(P_t f)\,d\mu = -\int \frac{1}{P_t f}|\nabla P_t f|^2\,d\mu, \tag{65}$$

$$P_\infty f \;:=\; \lim_{t\uparrow\infty} P_t f = \int f\,d\mu \;=\; 1. \tag{66}$$

Indeed, the left–hand side of (55) can be reformulated as

$$\begin{aligned}\int g\,f\,d\mu - \int g\,d\mu\int f\,d\mu &\overset{(62),(66)}{=} \int g\,(P_0 f - P_\infty f)\,d\mu\\ &= \int_0^\infty \frac{d}{dt}\int g\,P_t f\,d\mu\,dt\\ &\overset{(63)}{=} -\int_0^\infty\int \nabla g\cdot\nabla P_t f\,d\mu\,dt.\end{aligned}$$

This yields the estimate

$$
\begin{aligned}
&\left|\int g\,f\,d\mu - \int g\,d\mu \int f\,d\mu\right| \\
&\leq \sup_x |\nabla g| \int_0^\infty \int |\nabla P_t f|\,d\mu\,dt \\
&\leq \sup_x |\nabla g| \int_0^\infty \left(\int P_t f\,d\mu \int \frac{1}{P_t f}|\nabla P_t f|^2\,d\mu\right)^{1/2} dt \\
&\overset{(64)}{=} \sup_x |\nabla g| \int_0^\infty \left(\int \frac{1}{P_t f}|\nabla P_t f|^2\,d\mu\right)^{1/2} dt.
\end{aligned}
$$

It remains to estimate the last term:

$$
\begin{aligned}
&\int_0^\infty \left(\int \frac{1}{P_t f}|\nabla P_t f|^2\,d\mu\right)^{1/2} dt \\
&\quad = \int_0^\infty \left(\int \frac{1}{P_t f}|\nabla P_t f|^2\,d\mu\right)^{-1/2} \int \frac{1}{P_t f}|\nabla P_t f|^2\,d\mu\,dt \\
&\quad \overset{(54),(64)}{\leq} \left(\frac{1}{2\rho}\right)^{1/2} \int_0^\infty \left(\int \Phi(P_t f)\,d\mu\right)^{-1/2} \int \frac{1}{P_t f}|\nabla P_t f|^2\,d\mu\,dt \\
&\quad \overset{(65)}{=} -\left(\frac{1}{2\rho}\right)^{1/2} \int_0^\infty \left(\int \Phi(P_t f)\,d\mu\right)^{-1/2} \frac{d}{dt}\int \Phi(P_t f)\,d\mu\,dt \\
&\quad = -\left(\frac{2}{\rho}\right)^{1/2} \int_0^\infty \frac{d}{dt}\left(\int \Phi(P_t f)\,d\mu\right)^{1/2} dt \\
&\quad = \left(\frac{2}{\rho}\right)^{1/2} \left(\left(\int \Phi(P_0 f)\,d\mu\right)^{1/2} - \left(\int \Phi(P_\infty f)\,d\mu\right)^{1/2}\right) \\
&\quad \overset{(62),(66)}{=} \left(\frac{2}{\rho}\right)^{1/2} \left(\int \Phi(f)\,d\mu\right)^{1/2}.
\end{aligned}
$$

Proof of Corollary 1.

Let $g(x)$ and $h(x)$ be given. We may assume that $h$ is bounded so that for sufficiently small $\epsilon > 0$ we have

$$
f(x) \;:=\; 1 + \epsilon\, h(x) \;\geq\; 0.
$$

We then may apply Lemma 5 to $f$ and $g$, which yields

$$\begin{aligned}&\epsilon\left|\int g\,h\,d\mu-\int g\,d\mu\int h\,d\mu\right|\\&\le\ \frac{1}{\rho}\sup_x|\nabla g|\left(1+\epsilon\int h\,d\mu\right)^{1/2}\left(\epsilon^2\int\frac{1}{1+\epsilon\,h}|\nabla h|^2\,d\mu\right)^{1/2}.\end{aligned}$$

Dividing by $\epsilon$ and letting it tend to zero yields

$$\left|\int g\,h\,d\mu-\int g\,d\mu\int h\,d\mu\right|\ \le\ \frac{1}{\rho}\sup_x|\nabla g|\left(\int|\nabla h|^2\,d\mu\right)^{1/2},$$

which is a stronger version of (56).

Proof of Lemma 6.

From the representation

$$\bar f(x_1)\ =\ \int f(x_1,x_2)\,\mu(dx_2|x_1)\ =\ \frac{\int f(x_1,x_2)\ \exp(-H(x_1,x_2))\,dx_2}{\int\exp(-H(x_1,x_2))\,dx_2},$$

we deduce the formula

$$\begin{aligned}\nabla_1\bar f(x_1)=&\frac{\int\nabla_1 f(x_1,x_2)\ \exp(-H(x_1,x_2))\,dx_2}{\int\exp(-H(x_1,x_2))\,dx_2}\\&-\frac{\int f(x_1,x_2)\,\nabla_1 H(x_1,x_2)\ \exp(-H(x_1,x_2))\,dx_2}{\int\exp(-H(x_1,x_2))\,dx_2}\\&+\frac{\int f(x_1,x_2)\ \exp(-H(x_1,x_2))\,dx_2}{\int\exp(-H(x_1,x_2))\,dx_2}\\&\times\frac{\int\nabla_1 H(x_1,x_2)\ \exp(-H(x_1,x_2))\,dx_2}{\int\exp(-H(x_1,x_2))\,dx_2}\\=&\int\nabla_1 f(x_1,x_2)\,\mu(dx_2|x_1)\\&-\left(\int f(x_1,x_2)\,\nabla_1 H(x_1,x_2)\,\mu(dx_2|x_1)\right.\\&\left.-\int f(x_1,x_2)\,\mu(dx_2|x_1)\int\nabla_1 H(x_1,x_2)\,\mu(dx_2|x_1)\right).\end{aligned}$$

Hence Lemma 5, applied to $\mu(dx_2|x_1)$, $f(x_1,x_2)$, and $g(x_2)=\nabla_1 H(x_1,x_2)$ for fixed $x_1$, yields

$$
\begin{aligned}
|\nabla_1 \bar{f}(x_1)| \leq & \int |\nabla_1 f(x_1,x_2)|\,\mu(dx_2|x_1) \\
& + \frac{1}{\rho_2} \sup_{x_2} |\nabla_2\nabla_1 H(x_1,x_2)| \left(\int f(x_1,x_2)\,\mu(dx_2|x_1)\right)^{1/2} \\
& \left(\int \frac{1}{f(x_1,x_2)}|\nabla_2 f(x_1,x_2)|^2\,\mu(dx_2|x_1)\right)^{1/2} \\
\leq & (\bar{f}(x_1))^{1/2} \left( \left(\int \frac{1}{f(x_1,x_2)}|\nabla_1 f(x_1,x_2)|^2\,\mu(dx_2|x_1)\right)^{1/2} \right. \\
& \left. + \frac{\kappa_{12}}{\rho_2} \left(\int \frac{1}{f(x_1,x_2)}|\nabla_2 f(x_1,x_2)|^2\,\mu(dx_2|x_1)\right)^{1/2} \right).
\end{aligned}
$$

We rewrite this inequality as

$$
\begin{aligned}
& \left(\frac{1}{\bar{f}(x_1)}|\nabla_1 \bar{f}(x_1)|^2\right)^{1/2} \\
\leq & \left(\int \frac{1}{f(x_1,x_2)}|\nabla_1 f(x_1,x_2)|^2\,\mu(dx_2|x_1)\right)^{1/2} \\
& + \frac{\kappa_{12}}{\rho_2} \left(\int \frac{1}{f(x_1,x_2)}|\nabla_2 f(x_1,x_2)|^2\,\mu(dx_2|x_1)\right)^{1/2}.
\end{aligned}
$$

The triangle inequality in $L^2(\bar{\mu}(dx_1))$ yields the desired result.

PROOF OF LEMMA 7.

Let an arbitrary $f(x_2) \geq 0$ be given. We set for abbreviation

$$
\bar{f}(x_1) = \int f(x_2)\,\mu(dx_2|x_1). \tag{67}
$$

We split the left–hand side of the LSI as follows:

$$
\begin{aligned}
&\int \Phi(f(x_2))\bar{\mu}(dx_2) - \Phi\left(\int f(x_2)\,\bar{\mu}(dx_2)\right) \\
&= \int\int \Phi(f(x_2))\,\mu(dx_2|x_1)\,\bar{\mu}(dx_1) - \Phi\left(\int\int f(x_2)\,\mu(dx_2|x_1)\,\bar{\mu}(dx_1)\right) \\
&\overset{(67)}{=} \int \left(\int \Phi(f(x_2))\,\mu(dx_2|x_1) - \Phi\left(\int f(x_2)\,\mu(dx_2|x_1)\right)\right)\,\bar{\mu}(dx_1) \\
&\quad + \int \Phi(\bar{f}(x_1))\,\bar{\mu}(dx_1) - \Phi\left(\int \bar{f}(x_1)\,\bar{\mu}(dx_1)\right).
\end{aligned}
$$

According to our assumptions (58) and (59), this yields the estimate

$$
\begin{aligned}
&\left|\int \Phi(f(x_2))\bar{\mu}(dx_2) - \Phi\left(\int f(x_2)\,\bar{\mu}(dx_2)\right)\right| \\
&\leq \int \frac{1}{\rho_2} \int \frac{1}{2f(x_2)}|\nabla_2 f(x_2)|^2\,\mu(dx_2|x_1)\,\bar{\mu}(dx_1) \\
&\quad + \frac{1}{\bar{\rho}_1}\int \frac{1}{2\bar{f}(x_1)}|\nabla_1 \bar{f}(x_1)|^2\,\bar{\mu}(dx_1) \\
&= \frac{1}{\rho_2}\int \frac{1}{2f}|\nabla_2 f|^2\,\bar{\mu}(dx_2) + \frac{1}{\bar{\rho}_1}\int \frac{1}{2\bar{f}}|\nabla_1 \bar{f}|^2\,\bar{\mu}(dx_1). \qquad (68)
\end{aligned}
$$

We now apply Lemma 6 to the last term. Since $f$ does not depend on $x_1$, this yields

$$
\begin{aligned}
\int \frac{1}{2\bar{f}}|\nabla_1 \bar{f}|^2\,\bar{\mu}(dx_1) &\leq \frac{\kappa_{12}^2}{\rho_2^2}\int \frac{1}{2f}|\nabla_2 f|^2\,d\mu \\
&= \frac{\kappa_{12}^2}{\rho_2^2}\int \frac{1}{2f}|\nabla_2 f|^2\,\bar{\mu}(dx_2). \qquad (69)
\end{aligned}
$$

Inserting (69) into (68) yields as desired

$$
\begin{aligned}
&\left|\int \Phi(f(x_2))\bar{\mu}(dx_2) - \Phi\left(\int f(x_2)\,\bar{\mu}(dx_2)\right)\right| \\
&\qquad\qquad \leq \left(\frac{1}{\rho_2} + \frac{1}{\bar{\rho}_1}\frac{\kappa_{12}^2}{\rho_2^2}\right)\int \frac{1}{2f}|\nabla_2 f|^2\,\bar{\mu}(dx_2).
\end{aligned}
$$

PROOF OF COROLLARY 2.

By an approximation argument, we may assume that we have for the marginals

$$\bar{\mu}(dx_1) \quad \text{satisfies LSI}(\bar{\rho}_1),$$

$$\bar{\mu}(dx_2) \quad \text{satisfies LSI}(\bar{\rho}_2),$$

for some constants $\bar{\rho}_1, \bar{\rho}_2 > 0$. Lemma 7 now yields

$$\frac{1}{\bar{\rho}_2} \ \leq \ \frac{1}{\rho_2} + \frac{1}{\bar{\rho}_1}\,\frac{\kappa_{12}^2}{\rho_2^2}. \tag{70}$$

By symmetry, we also have

$$\frac{1}{\bar{\rho}_1} \ \leq \ \frac{1}{\rho_1} + \frac{1}{\bar{\rho}_2}\,\frac{\kappa_{12}^2}{\rho_1^2}. \tag{71}$$

Inserting (70) into (71) yields

$$\frac{1}{\bar{\rho}_1} \ \leq \ \frac{1}{\rho_1} + \left(\frac{1}{\rho_2} + \frac{1}{\bar{\rho}_1}\,\frac{\kappa_{12}^2}{\rho_2^2}\right)\frac{\kappa_{12}^2}{\rho_1^2}$$

and thus

$$\left(1 - \frac{\kappa_{12}^4}{\rho_1^2\,\rho_2^2}\right)\frac{1}{\bar{\rho}_1} \ \leq \ \frac{1}{\rho_1} + \frac{1}{\rho_2}\,\frac{\kappa_{12}^2}{\rho_1^2},$$

which we rewrite as

$$\left(1 - \frac{\kappa_{12}^2}{\rho_1\,\rho_2}\right)\left(1 + \frac{\kappa_{12}^2}{\rho_1\,\rho_2}\right)\frac{1}{\bar{\rho}_1} \ \leq \ \left(1 + \frac{\kappa_{12}^2}{\rho_1\,\rho_2}\right)\frac{1}{\rho_1}.$$

This yields

$$\left(1 - \frac{\kappa_{12}^2}{\rho_1\,\rho_2}\right)\frac{1}{\bar{\rho}_1} \ \leq \ \frac{1}{\rho_1}$$

and thus

$$\bar{\rho}_1 \ \geq \ \left(1 - \frac{\kappa_{12}^2}{\rho_1\,\rho_2}\right)\rho_1 \ = \ \rho_1 - \frac{\kappa_{12}^2}{\rho_2},$$

which is positive by assumption.

PROOF OF LEMMA 8.

As the starting point we have the two formulas

$$\nabla_1 \bar{H}(x_1,x_2) \;=\; \frac{\int \nabla_1 H(x_1,x_2,x_3)\, \exp(-H(x_1,x_2,x_3))\,dx_3}{\int \exp(-H(x_1,x_2,x_3))\,dx_3}$$

and

$$\begin{aligned}
&\nabla_2\nabla_1 \bar{H}(x_1,x_2)\\
&= \frac{\int \nabla_2\nabla_1 H(x_1,x_2,x_3)\, \exp(-H(x_1,x_2,x_3))\,dx_3}{\int \exp(-H(x_1,x_2,x_3))\,dx_3}\\
&\quad - \frac{\int \nabla_1 H(x_1,x_2,x_3) \otimes \nabla_2 H(x_1,x_2,x_3)\, \exp(-H(x_1,x_2,x_3))\,dx_3}{\int \exp(-H(x_1,x_2,x_3))\,dx_3}\\
&\quad + \frac{\int \nabla_1 H(x_1,x_2,x_3)\, \exp(-H(x_1,x_2,x_3))\,dx_3}{\int \exp(-H(x_1,x_2,x_3))\,dx_3}\\
&\quad \otimes \frac{\int \nabla_2 H(x_1,x_2,x_3)\, \exp(-H(x_1,x_2,x_3))\,dx_3}{\int \exp(-H(x_1,x_2,x_3))\,dx_3}\\
&= \int \nabla_2\nabla_1 H(x_1,x_2,x_3)\, \mu(dx_3|x_1,x_2)\\
&\quad - \Big( \int \nabla_1 H(x_1,x_2,x_3) \otimes \nabla_2 H(x_1,x_2,x_3)\, \mu(dx_3|x_1,x_2)\\
&\quad - \int \nabla_1 H(x_1,x_2,x_3)\, \mu(dx_3|x_1,x_2) \otimes \int \nabla_2 H(x_1,x_2,x_3)\, \mu(dx_3|x_1,x_2) \Big).
\end{aligned}$$

According to Corollary 1 applied to $\mu(dx_3|x_1,x_2)$, $g(x_3) = \nabla_1 H(x_1,x_2,x_3)$, and $h(x_3) = \nabla_2 H(x_1,x_2,x_3)$, we have the inequality

$$\begin{aligned}
|\nabla_2\nabla_1 \bar{H}(x_1,x_2)| &\le \sup_{x_3} |\nabla_2\nabla_1 H(x_1,x_2,x_3)|\\
&\quad + \frac{1}{\rho_3} \sup_{x_3} |\nabla_3\nabla_1 H(x_1,x_2,x_3)| \sup_{x_3} |\nabla_3\nabla_1 H(x_1,x_2,x_3)|\\
&\le \kappa_{12} + \frac{1}{\rho_3}\, \kappa_{13}\, \kappa_{23}.
\end{aligned}$$

Taking the sup over $(x_1,x_2)$ yields (61).

PROOF OF LEMMA 9.

We prove Lemma 9 by induction in $N$. The case $N = 1$ is trivial. We now assume that the lemma holds for $N-1 \ge 1$ and argue that it also holds for

$N$. To this purpose, we introduce the related block partitioning of $A$:

$$A \ = \ \begin{pmatrix} A' & -\kappa \\ & \\ -\kappa^t & \rho \end{pmatrix}.$$

The inverse $A^{-1}$ is given by the block partitioning

$$A^{-1} \ = \ \begin{pmatrix} (A')^{-1} + \dfrac{(A')^{-1}\kappa \otimes (A')^{-1}\kappa}{\rho - \kappa \cdot (A')^{-1}\kappa} & \dfrac{(A')^{-1}\kappa}{\rho - \kappa \cdot (A')^{-1}\kappa} \\ & \\ \left(\dfrac{(A')^{-1}\kappa}{\rho - \kappa \cdot (A')^{-1}\kappa}\right)^t & \dfrac{1}{\rho - \kappa \cdot (A')^{-1}\kappa} \end{pmatrix}.$$

Based on this representation, we now argue that the entries of $A^{-1}$ are non–negative. As an immediate consequence of the positive definiteness of $A$ we have

$$\rho - \kappa \cdot (A')^{-1}\kappa \ > \ 0, \tag{72}$$

so that

$$(A^{-1})_{NN} \ = \ \frac{1}{\rho - \kappa \cdot (A')^{-1}\kappa} \ \geq \ 0.$$

Combining the induction hypothesis applied to $A'$, i.e.

$$((A')^{-1})_{ij} \ \geq \ 0 \quad \text{for } i, j \in \{1, \cdots, N-1\}, \tag{73}$$

with our assumption

$$\kappa_j \ \geq \ 0 \quad \text{for } j \in \{1, \cdots, N-1\}, \tag{74}$$

we obtain

$$((A')^{-1}\kappa)_i \ = \ \sum_j ((A')^{-1})_{ij}\kappa_j \ \geq \ 0 \quad \text{for } i \in \{1, \cdots, N-1\}, \tag{75}$$

Together with (72), we obtain

$$(A^{-1})_{iN} \ = \ \frac{((A')^{-1}\kappa)_i}{\rho - \kappa \cdot (A')^{-1}\kappa} \ \geq \ 0 \quad \text{for } i \in \{1, \cdots, N-1\}. \tag{76}$$

From (75) we obtain

$$((A')^{-1}\kappa\otimes(A')^{-1}\kappa)_{ij} \;=\; ((A')^{-1}\kappa)_i\,((A')^{-1}\kappa)_j \;\geq\; 0 \quad \text{for } i,j\in\{1,\cdots,N-1\}$$

and together with (72) and (73):

$$(A^{-1})_{ij} \;=\; ((A')^{-1})_{ij}+\frac{((A')^{-1}\kappa\otimes(A')^{-1}\kappa)_{ij}}{\rho-\kappa\cdot(A')^{-1}\kappa} \;\geq\; 0 \quad \text{for } i,j\in\{1,\cdots,N-1\}.$$

### *4.2 Proofs of Theorems 1 and 2*

PROOF OF THEOREM 1.

We shall prove the seemingly stronger result

$$\begin{aligned}&\forall\, f(x_1,\cdots,x_N)\geq 0 \quad \int\Phi(f)\,d\mu-\Phi\left(\int f d\mu\right)\\ &\leq \sum_{i,j\in\{1,\cdots,N\}}(A^{-1})_{ij}\left(\int\frac{1}{2f}|\nabla_i f|^2\,d\mu\right)^{1/2}\left(\int\frac{1}{2f}|\nabla_j f|^2\,d\mu\right)^{1/2},\end{aligned} \tag{77}$$

where $(A^{-1})_{ij}$ denote the coefficients of the inverse $A^{-1}$ of $A$. Statement (77) indeed implies (6): According to (5) we have $A^{-1}\leq\frac{1}{\rho}\,\mathrm{id}$ in the sense of quadratic forms so that (77) implies

$$\begin{aligned}&\int\Phi(f)\,d\mu-\Phi\left(\int f d\mu\right)\\ &\leq \sum_{i,j\in\{1,\cdots,N\}}\frac{1}{\rho}\,\delta_{ij}\left(\int\frac{1}{2f}|\nabla_i f|^2\,d\mu\right)^{1/2}\left(\int\frac{1}{2f}|\nabla_j f|^2\,d\mu\right)^{1/2}\\ &=\frac{1}{\rho}\int\frac{1}{2f}\sum_{i\in\{1,\cdots,N\}}|\nabla_i f|^2\,d\mu.\end{aligned}$$

We show (77) by induction in $N$. For $N=1$, the statement (77) is a trivial consequence of our assumption (3). We thus assume that we know (77) for any

$(N-1)$–component system and argue that it holds for $N$. It will be convenient to work with the related block decomposition of $A$:

$$A \;=\; \begin{pmatrix} A' & -\kappa_N \\ & \\ -\kappa_N^t & \rho_N \end{pmatrix}. \tag{78}$$

Denote by $\bar{A}$ the $(N-1)\times(N-1)$–matrix defined by

$$\bar{A} \;=\; A' - \frac{1}{\rho_N}\,\kappa_N \otimes \kappa_N. \tag{79}$$

We observe that $\bar{A}$ inherits our assumptions on $A$: It is symmetric and positive definite.

We start by considering the system $\bar{\mu}(dx_1 \cdots dx_{N-1})$, i.e. the marginal of $\mu(dx_1 \cdots dx_N)$ on $X_1 \times \cdots \times X_{N-1}$. Its Hamiltonian is given by

$$\bar{H}(x_1, \cdots, x_{N-1}) \;=\; -\log \int \exp(-H(x_1, \cdots, x_{N-1}, x_N))\, dx_N.$$

Let $i < j \in \{1, \cdots, N-1\}$ be arbitrary. Lemma 8 applied to $\mu(dx_i dx_j dx_N | \cdots)$ yields

$$\forall\ (x_1, \cdots, x_{N-1}) \quad |\nabla_i \nabla_j \bar{H}(x_1, \cdots, x_{N-1})| \;\le\; \bar{\kappa}_{ij}$$

with

$$\bar{\kappa}_{ij} \;\le\; \kappa_{ij} + \frac{\kappa_{iN}\,\kappa_{jN}}{\rho_N} \;\overset{(79)}{=}\; -\bar{A}_{ij}.$$

Now let $i \in \{1, \cdots, N-1\}$ be arbitrary. Corollary 2 applied to $\mu(dx_i dx_N | \cdots)$ yields

$$\begin{aligned} &\forall\ (x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_{N-1}) \\ &\bar{\mu}(dx_i | x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_{N-1}) \quad \text{satisfies LSI}(\bar{\rho}_i) \end{aligned} \tag{80}$$

with

$$\bar\rho_i \;\ge\; \rho_i - \frac{\kappa_{iN}^2}{\rho_N} \;\overset{(79)}{=}\; \bar A_{ii}.$$

Thus, we may apply the induction hypothesis to $\bar\mu(dx_1\cdots dx_{N-1})$ and $\bar A$:

$$\begin{aligned}
&\forall\ \bar f(x_1,\cdots,x_{N-1})\ge 0 \quad \int\Phi(\bar f)\,d\bar\mu-\Phi\left(\int \bar f d\bar\mu\right)\\
&\le \sum_{i,j\in\{1,\cdots,N-1\}}(\bar A^{-1})_{ij}\left(\int\frac{1}{2\bar f}|\nabla_i\bar f|^2\,d\bar\mu\right)^{1/2}\left(\int\frac{1}{2\bar f}|\nabla_j\bar f|^2\,d\bar\mu\right)^{1/2}. \qquad (81)
\end{aligned}$$

Now let $f(x_1,\cdots,x_N)\ge 0$ be given and set

$$\bar f(x_1,\cdots,x_{N-1}) \;:=\; \int f(x_1,\ldots,x_{N-1},x_N)\,\mu(dx_N|x_1,\cdots,x_{N-1}).$$

As in the proof of Lemma 7, we split the left–hand side of (77):

$$\begin{aligned}
&\int\Phi(f)\,d\mu-\Phi\left(\int f\,d\mu\right)\\
&=\int\left(\int\Phi(f)\,\mu(dx_N|\cdots)-\Phi\left(\int f\,\mu(dx_N|\cdots)\right)\right)\\
&\qquad \bar\mu(dx_1\cdots dx_{N-1})\\
&\quad+\int\Phi(\bar f)\,d\bar\mu-\Phi\left(\int\bar f\,d\bar\mu\right). \qquad (82)
\end{aligned}$$

By assumption (3) we have

$$\begin{aligned}
&\int\Phi(f)\,\mu(dx_N|\cdots)-\Phi\left(\int f\,\mu(dx_N|\cdots)\right)\\
&\qquad\le\frac{1}{\rho_N}\int\frac{1}{2f}|\nabla_N f|^2\,\mu(dx_N|\cdots),
\end{aligned}$$

so that we obtain for the first right–hand side term in (82):

$$\begin{aligned}
&\int\left(\int\Phi(f)\,\mu(dx_N|\cdots)-\Phi\left(\int f\,\mu(dx_N|\cdots)\right)\right)\bar\mu(dx_1\cdots dx_{N-1})\\
&\le\frac{1}{\rho_N}\int\frac{1}{2f}|\nabla_N f|^2\,d\mu. \qquad (83)
\end{aligned}$$

We apply (81) to the second right–hand side term in (82). Combining this with (83), (82) becomes

$$\begin{aligned}
&\int \Phi(f)\,d\mu - \Phi\left(\int f d\mu\right) \\
&\leq \sum_{i,j\in\{1,\cdots,N-1\}} (\bar{A}^{-1})_{ij} \left(\int \frac{1}{2\bar{f}}|\nabla_i \bar{f}|^2\,d\bar{\mu}\right)^{1/2} \left(\int \frac{1}{2\bar{f}}|\nabla_j \bar{f}|^2\,d\bar{\mu}\right)^{1/2} \\
&\quad + \frac{1}{\rho_N}\int \frac{1}{2f}|\nabla_N f|^2\,d\mu. \qquad (84)
\end{aligned}$$

We now want to express the first right–hand side terms of (84) in terms of $f$ and $\mu$. To this purpose, let $i \in \{1,\cdots,N-1\}$ be arbitrary. Because of our assumptions (2) and (3), we may apply Lemma 6 to $\mu(dx_i dx_N|\cdots)$ and obtain

$$\begin{aligned}
&\left(\int \frac{1}{2\bar{f}}|\nabla_i \bar{f}|^2\,\bar{\mu}(dx_i|\cdots)\right)^{1/2} \\
&\leq \left(\int \frac{1}{2f}|\nabla_i f|^2\,\mu(dx_i dx_N|\cdots)\right)^{1/2} + \frac{\kappa_{iN}}{\rho_N}\left(\int \frac{1}{2f}|\nabla_N f|^2\,\mu(dx_i dx_N|\cdots)\right)^{1/2}.
\end{aligned}$$

By the triangle inequality in $L^2(\bar{\mu}(dx_1\cdots dx_{i-1}dx_{i+1}\cdots dx_{N-1}))$, this yields

$$\begin{aligned}
&\left(\int \frac{1}{2\bar{f}}|\nabla_i \bar{f}|^2\,d\bar{\mu}\right)^{1/2} \\
&\leq \left(\int \frac{1}{2f}|\nabla_i f|^2\,d\mu\right)^{1/2} + \frac{\kappa_{iN}}{\rho_N}\left(\int \frac{1}{2f}|\nabla_N f|^2\,d\mu\right)^{1/2}. \qquad (85)
\end{aligned}$$

Since $\bar{A}$ has non–positive off–diagonal entries, an application of Lemma 9 yields that all entries of $(\bar{A})^{-1}$ are non–negative. Thus we may insert the inequality (85) into (84):

$$
\begin{aligned}
&\int \Phi(f)\,d\mu - \Phi\left(\int f d\mu\right) \\
&\leq \sum_{i,j\in\{1,\cdots,N-1\}} (\bar{A}^{-1})_{ij} \left(\int \frac{1}{2f}|\nabla_i f|^2\,d\mu\right)^{1/2} \left(\int \frac{1}{2f}|\nabla_j f|^2\,d\mu\right)^{1/2} \\
&\quad +2 \sum_{i,j\in\{1,\cdots,N-1\}} (\bar{A}^{-1})_{ij}\,\frac{\kappa_{jN}}{\rho_N} \left(\int \frac{1}{2f}|\nabla_i f|^2\,d\mu\right)^{1/2} \left(\int \frac{1}{2f}|\nabla_N f|^2\,d\mu\right)^{1/2} \\
&\quad + \left(\frac{1}{\rho_N} + \sum_{i,j\in\{1,\cdots,N-1\}} \frac{\kappa_{iN}}{\rho_N}\,(\bar{A}^{-1})_{ij}\,\frac{\kappa_{jN}}{\rho_N}\right) \int \frac{1}{2f}|\nabla_N f|^2\,d\mu. \qquad (86)
\end{aligned}
$$

We now argue that (86) and (77) coincide, which amounts to argue that the matrix

$$
\begin{pmatrix}
\bar{A}^{-1} & \dfrac{\bar{A}^{-1}\kappa_N}{\rho_N} \\
\left(\dfrac{\bar{A}^{-1}\kappa_N}{\rho_N}\right)^t & \dfrac{1}{\rho_N} + \dfrac{\kappa_N\cdot\bar{A}^{-1}\kappa_N}{\rho_N^2}
\end{pmatrix}
$$

is the inverse of

$$
A \;=\; \begin{pmatrix}
A' & -\kappa_N \\
-\kappa_N^t & \rho_N
\end{pmatrix}
\overset{(79)}{=}
\begin{pmatrix}
\bar{A} + \dfrac{\kappa_N\otimes\kappa_N}{\rho_N} & -\kappa_N \\
-\kappa_N^t & \rho_N
\end{pmatrix}.
$$

This reduces to the four obvious identities:

$$
\begin{aligned}
\bar{A}^{-1}\left(\bar{A} + \frac{\kappa_N\otimes\kappa_N}{\rho_N}\right) + \frac{\bar{A}^{-1}\kappa_N}{\rho_N}\otimes(-\kappa_N) &= \mathrm{id}', \\
\bar{A}^{-1}(-\kappa_N) + \rho_N\frac{\bar{A}^{-1}\kappa_N}{\rho_N} &= 0 \\
\left(\bar{A} + \frac{\kappa_N\otimes\kappa_N}{\rho_N}\right)\frac{\bar{A}^{-1}\kappa_N}{\rho_N} - \left(\frac{1}{\rho_N} + \frac{\kappa_N\cdot\bar{A}^{-1}\kappa_N}{\rho_N^2}\right)(\kappa_N) &= 0 \\
\frac{\bar{A}^{-1}\kappa_N}{\rho_N}\cdot(-\kappa_N) + \left(\frac{1}{\rho_N} + \frac{\kappa_N\cdot\bar{A}'\kappa_N}{\rho_N^2}\right)\rho_N &= 1.
\end{aligned}
$$

PROOF OF THEOREM 2.

Recalling the definition of $\bar f$ from (57), we integrate the identity

$$\Phi(f) \;=\; f\,\log\bar f \,+\, \bar f\,\Phi\Big(\frac{f}{\bar f}\Big)$$

and use Definition 2 to obtain the usual decomposition of entropy:

$$\begin{aligned}
&\int\Phi(f)d\mu-\Phi\left(\int fd\mu\right)\\
&=\int\int f\,\log\bar f\,\mu(dx_2|x_1)\,\bar\mu(dx_1)+\int\int\bar f\,\Phi\Big(\frac{f}{\bar f}\Big)\,\mu(dx_2|x_1)\,\bar\mu(dx_1)\\
&\qquad-\Phi\left(\int fd\mu\right)\\
&=\int\int f\,\mu(dx_2|x_1)\,\log\bar f\,\bar\mu(dx_1)+\int\bar f\,\int\Phi\Big(\frac{f}{\bar f}\Big)\,\mu(dx_2|x_1)\,\bar\mu(dx_1)\\
&\qquad-\Phi\left(\int fd\mu\right)\\
&\overset{(57)}{=}\;\int\Phi(\bar f)\,\bar\mu(dx_1)-\Phi\left(\int\bar f\bar\mu(dx_1)\right)\\
&\qquad\qquad+\int\bar f\,\int\Phi\Big(\frac{f}{\bar f}\Big)\,\mu(dx_2|x_1)\bar\mu(dx_1). \qquad (87)
\end{aligned}$$

We start with the last term on the right–hand side of (87). According to (23), we have for all $x_1\in X_1$

$$\begin{aligned}
\int\Phi\Big(\frac{f}{\bar f}\Big)\,\mu(dx_2|x_1)\;&\le\;\Phi\Big(\int\frac{f}{\bar f}\,\mu(dx_2|x_1)\Big)+\frac{1}{\rho_2}\int\frac{\bar f}{2f}\,\Big|\nabla_2\frac{f}{\bar f}\Big|^2\,\mu(dx_2|x_1)\\
&\overset{(32)}{=}\;\frac{1}{\rho_2}\,\frac{1}{\bar f}\int\frac{1}{2f}\,|\nabla_2 f|^2\,\mu(dx_2|x_1),
\end{aligned}$$

so that by integrating we obtain

$$\begin{aligned}
\int\bar f\int\Phi\Big(\frac{f}{\bar f}\Big)\,\mu(dx_2|x_1)\,\bar\mu(dx_1)\;&\le\;\frac{1}{\rho_2}\int\int\frac{1}{2f}|\nabla_2 f|^2\,\mu(dx_2|x_1)\,\bar\mu(dx_1)\\
&=\;\frac{1}{\rho_2}\int\frac{1}{2f}|\nabla_2 f|^2\,d\mu. \qquad (88)
\end{aligned}$$

Turning to the first two terms on the right–hand side of (87), we observe:

$$\int\Phi(\bar f)\,\bar\mu(dx_1)-\Phi\Big(\int\bar f\,\bar\mu(dx_1)\Big)\overset{(24)}{\le}\;\frac{1}{\bar\rho_1}\;\int\frac{1}{2\bar f}|\nabla_1\bar f|^2\,\bar\mu(dx_1). \qquad (89)$$

By Lemma 6 and Young's inequality, we have for any $\tau \in (0,1)$,

$$\begin{aligned} &\int \frac{1}{2\bar{f}} |\nabla_1 \bar{f}|^2 \, \bar{\mu}(dx_1) \\ &\leq \frac{1}{\tau} \int \frac{1}{2f} |\nabla_1 f|^2 \, d\mu \, + \, \frac{1}{1-\tau} \frac{\kappa_{12}^2}{\rho_2^2} \int \frac{1}{2f} |\nabla_2 f|^2 \, d\mu, \end{aligned}$$

so that (89) becomes

$$\begin{aligned} &\int \Phi(\bar{f}) \, \bar{\mu}(dx_1) - \Phi\Big( \int f \, d\mu \Big) \\ &\leq \frac{1}{\tau \bar{\rho}_1} \int \frac{1}{2f} |\nabla_1 f|^2 \, d\mu \, + \, \frac{1}{1-\tau} \frac{\kappa_{12}^2}{\rho_2^2 \, \bar{\rho}_1} \int \frac{1}{2f} |\nabla_2 f|^2 \, d\mu. \end{aligned} \tag{90}$$

Substituting (88) and (90) into (87) gives:

$$\begin{aligned} &\int \Phi(f) d\mu - \Phi\Big( \int f \, d\mu \Big) \\ &\leq \frac{1}{\rho_2} \int \frac{1}{2f} |\nabla_2 f|^2 \, d\mu \\ &\qquad + \frac{1}{\tau} \frac{1}{\bar{\rho}_1} \int \frac{1}{2f} |\nabla_1 f|^2 \, d\mu + \frac{1}{1-\tau} \frac{\kappa_{12}^2}{\rho_2^2 \, \bar{\rho}_1} \int \frac{1}{2f} |\nabla_2 f|^2 d\mu. \end{aligned}$$

Since $|\nabla f|^2 = |\nabla_1 f|^2 + |\nabla_2 f|^2$, this yields the bound on the LSI constant:

$$\frac{1}{\rho} \;\leq\; \max\left\{ \frac{1}{\tau} \frac{1}{\bar{\rho}_1}, \frac{1}{\rho_2} + \frac{1}{1-\tau} \frac{\kappa_{12}^2}{\rho_2^2 \, \bar{\rho}_1} \right\}. \tag{91}$$

The optimization in $\tau$ completes the proof. Indeed, the optimal $\tau$ in (91) is characterized by

$$\frac{1}{\tau \bar{\rho}_1} \;=\; \frac{1}{\rho_2} + \frac{\kappa_{12}^2}{(1-\tau)\rho_2^2 \bar{\rho}_1},$$

that is,

$$(1-\tau)\rho_2^2 \;=\; \bar{\rho}_1 \rho_2 \tau (1-\tau) + \kappa_{12}^2 \tau.$$

The admissible solution is

$$\tau \;=\; \frac{1}{2} \left( \left( 1 + \frac{\kappa_{12}^2}{\rho_2 \bar{\rho}_1} \right) + \frac{\rho_2}{\bar{\rho}_1} - \left( \left( 1 + \frac{\kappa_{12}^2}{\rho_2 \bar{\rho}_1} + \frac{\rho_2}{\bar{\rho}_1} \right)^2 - 4 \frac{\rho_2}{\bar{\rho}_1} \right)^{1/2} \right),$$

so that (91) turns as desired into

$$\rho \geq \frac{1}{2}\left(\rho_2 + \bar{\rho}_1 + \frac{\kappa_{12}^2}{\rho_2} - \left(\left(\rho_2 + \bar{\rho}_1 + \frac{\kappa_{12}^2}{\bar{\rho}_1}\right)^2 - 4\rho_2\,\bar{\rho}_1\right)^{1/2}\right).$$

## Acknowledgements

We thank G. Blower, F. Bolley, A. Eberle, M. Rost, and H. T. Yau for inspiring discussions. We also gratefully acknowledge N. Grunewald and C. Villani, our co-authors on [GORV]. F. Otto was partially supported by the DFG through SFB 611. M. G. Reznikoff was partially supported by an NSF Mathematical Sciences Postdoctoral Research Fellowship.

## References

[BE] D. Bakry and M. Emery, Diffusions Hypercontractives, *Sem. Probab. XIX, Lecture Notes in Math.*, **1123**, Springer–Verlag, (1985), pp. 177–206.

[BB] G. Blower and F. Bolley, Concentration of Measure on Product Spaces with Applications to Markov Processes, *Studia Math.*, to appear.

[BH] T. Bodineau and B. Hellfer, The Log–Sobolev Inequality for Unbounded Spin Systems, *Journal of Functional Analysis*, **166**, (1999), pp. 168–178.

[G] L. Gross, Logarithmic Sobolev Inequalities, *Amer. J. Math.*, **97**, (1975), pp. 1061–1083.

[GORV] N. Grunewald, F. Otto, M. G. Reznikoff, and C. Villani, A Two–Scale Proof of a Logarithmic Sobolev Inequality, SFB preprint 2005.

[GZ] A. Guionnet and B. Zegarlinski, Lectures on Logarithmic Sobolev Inequalities, *Séminaire de Probabilités, XXXVI, Lecture Notes in Math.*, **1801**, Springer–Verlag, (2003), pp. 1–134.

[HS] R. Holley and D.Stroock, Logarithmic Sobolev Inequalities and Stochastic Ising Models, *J. Statist. Phys.*, **46**, (1987), pp. 1159–1194.

[L] M. Ledoux, Logarithmic Sobolev Inequalities for Unbounded Spin Systems Revisited, *Séminaire de Probabilités, XXXV, Lecture Notes in Math.*, **1755**, Springer–Verlag (2001), pp. 167–194.

[LY] S.L. Lu and H.T. Yau, Spectral Gap and Logarithmic Sobolev Inequality for Kawasaki and Glauber Dynamics, *Commun. Math. Phys.*, **156**, (1993), pp. 399–433.

[MO1] F. Martinelli and E. Olivieri, Approach to Equilibrium of Glauber Dynamics in the One–Phase Region. I. The Attractive Case, *Comm. Math. Phys.*, **161**, (1994), pp. 447–486.

[MO2] F. Martinelli and E. Olivieri, Approach to Equilibrium of Glauber Dynamics in the One–Phase Region. II. The General Case, *Comm. Math. Phys.*, **161**, (1994), pp. 487–514.

[O] F. Otto, The Geometry of Dissipative Evolution Equations: The Porous Medium Equation, *Commun. Partial Differential Equations*, **26**, (2001), pp. 101–174.

[OV] F. Otto and C. Villani, Generalization of an Inequality by Talagrand and Links with the Logarithmic Sobolev Inequality, *J. Funct. Analysis*, **173**, (2000), pp. 361–400.

[R] G. Royer, Une Initiation aux Inégalités de Sobolev Logarithmiques, Cours Spécialisés, Soc. Math. de France, (1999).

[SZ] D. W. Stroock and B. Zegarlinski, The Logarithmic Sobolev Inequality for Continuous Spin Systems on a Lattice, *J. Func. Anal.*, **104**, (1992), pp. 299–326.

[SZ2] D. W. Stroock and B. Zegarlinski, The Logarithmic Sobolev Inequality for Discrete Spin Systems on a Lattice, *Comm. Math. Phys.*, **149**, (1992), pp. 175–193.

[Yo1] N. Yoshida, The Log–Sobolev Inequality for Weakly Coupled Lattice Fields, *Probab. Theor. Relat. Fields*, **115**, (1999), pp. 1–40.

[Yo2] N. Yoshida, The Equivalence of the Log–Sobolev Inequality and a Mixing Condition for Unbounded Spin Systems on the Lattice, *Ann. Inst. H. Poincaré Probab. Statist.*, **37**, (2001), pp. 223–243.

[Z2] B. Zegarlinski, Log–Sobolev Inequalities for Infinite One Dimensional Lattice Systems, *Commun. Math. Phys.*, **133**, (1990), pp. 147–162.

[Z1] B. Zegarlinski, On Log–Sobolev Inequalities for Infinite Lattice Systems, *Lett. Math. Phys.*, **20**, (1990), pp. 173–182.

[Z3] B. Zegarlinski, The Strong Decay to Equilibrium for the Stochastic Dynamics of Unbounded Spin Systems on a Lattice, *Comm. Math. Phys.*, **175**, (1996), pp. 401–432.